DOCUMENT RESUME

ED 123 893 FL 007 674

AUTHOR	Campeau, Peggie L.; And Others
TITLE	The Identification and Description of Exemplary Bilingual Education Programs.
INSTITUTION	American Institutes for Research in the Behavioral Sciences, Palo Alto, Calif.
SPONS AGENCY	Office of Education (DHEW), Washington, D.C. Office of Planning, Budgeting, and Evaluation.
REPORT NO	AIR-48300-8/75-FR(I)
PUB DATE	Aug 75
CONTRACT	OEC-0-74-9331
NOTE	175p.; For related documents, see FL 007 675; pages that appear to be missing were blank and unnumbered on the original document.
EDRS PRICE	MF-$0.83 HC-$8.69 Plus Postage.
DESCRIPTORS	Biculturalism; *Bilingual Education; *Demonstration Programs; Educational Programs; Elementary Secondary Education; Evaluation Methods; Language Programs; *Program Descriptions; *Program Evaluation
IDENTIFIERS	*Elementary Secondary Education Act Title VII; ESEA Title VII; Maine; Texas

ABSTRACT

This is the final report of one of three studies in an overall project entitled "Evaluation of Bilingual Education Programs." The objectives of this study were to: (1) identify and recommend as exemplary up to 10 bilingual education programs for submission to the Dissemination Review Panel (DRP) of the Education Division, DHEW, and (2) develop detailed descriptions of those programs approved by the DRP for dissemination. Screenings and site visits reduced the candidate pool of 175 programs to 7. The DRP approved the following four for dissemination: Alice Independent School District, Alice, Texas; Aprendemos en Dos Idiomas, Corpus Christi, Texas; Bilingual Education Program, Houston, Texas; and St. John Valley Bilingual Education Program, Madawaska, Maine. Descriptions were written for each of the four in sufficient detail to provide local educators with ideas to guide them in implementing similar practices. Each program was described as it operated during 1973-74, but some program components were treated historically to illustrate major changes. The context within which each program developed and operated, and the special educational needs of the students, were also discussed. In addition, specific recommendations were made on ways in which local evaluations could be improved to increase chances for identifying exemplary programs. (Author/RM)

Documents acquired by ERIC include many informal unpublished materials not available from other sources. ERIC makes every effort to obtain the best copy available. Nevertheless, items of marginal reproducibility are often encountered and this affects the quality of the microfiche and hardcopy reproductions ERIC makes available via the ERIC Document Reproduction Service (EDRS). EDRS is not responsible for the quality of the original document. Reproductions supplied by EDRS are the best that can be made from the original.

AIR-48300-8/75-FR(I)

FL

ED123893

THE IDENTIFICATION AND DESCRIPTION OF EXEMPLARY BILINGUAL EDUCATION PROGRAMS

A Study Completed Under

Contract No. OEC 0-74-9331

Evaluation of Bilingual Education Programs

Peggie L. Campeau
A. Oscar H. Roberts
John E. Bowers
Melanie Austin
Sarah J. Roberts

U S DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

August 1975

FL007674

AMERICAN INSTITUTES FOR RESEARCH
Post Office Box 1113 / Palo Alto, California 94302

AIR-48300-8/75-FR(I)

THE IDENTIFICATION AND DESCRIPTION OF EXEMPLARY BILINGUAL EDUCATION PROGRAMS

A Study Completed Under
Contract No. OEC 0-74-9331
Evaluation of Bilingual Education Programs

Peggie L. Campeau
A. Oscar H. Roberts
John E. Bowers
Melanie Austin
Sarah J. Roberts

August 1975

AIR-48300-8/75-FR(I)

THE IDENTIFICATION AND DESCRIPTION OF EXEMPLARY BILINGUAL EDUCATION PROGRAMS

A Study Completed Under
Contract No. OEC-0-74-9331
Evaluation of Bilingual Education Programs

Peggie L. Campeau
A. Oscar H. Roberts
John E. Bowers
Melanie Austin
Sarah J. Roberts

American Institutes for Research
in the Behavioral Sciences
Palo Alto, California

August 1975

The research reported herein was performed pursuant to a contract with the Office of Education, U.S. Department of Health, Education, and Welfare. Contractors undertaking such projects under Government sponsorship are encouraged to express freely their professional judgment in the conduct of the project. Points of view or opinions stated do not, therefore, necessarily represent official Office of Education position or policy.

U. S. DEPARTMENT OF HEALTH, EDUCATION, AND WELFARE
Office of Education
Office of Planning, Budgeting, and Evaluation

RESOURCE APPENDIX

PREFACE

This document is the final report of one of three studies conducted by the American Institutes for Research (AIR) for the Office of Education, Office of Planning, Budgeting, and Evaluation, under Contract No. OEC-0-74-9331. The OE Project Officer was Edward B. Glassman. The study reported in this document was co-directed by Peggie Campeau and John Bowers, and the AIR project director for the overall contract was Malcolm N. Danoff.

The study reported here is entitled "The Identification and Description of Exemplary Bilingual Education Programs." Its objectives were to recommend to the Joint Dissemination Review Panel (DRP) of the Education Division, Department of Health, Education, and Welfare, those bilingual education programs with sound evidence demonstrating significantly improved student outcomes, and to prepare detailed descriptions for the programs that the DRP approved for dissemination.

ACKNOWLEDGEMENTS

We are grateful for the cooperation of program directors contacted throughout the study who generously gave their time to answer questions from AIR staff. Administrators and teachers at all the site-visited programs extended a special kind of hospitality to us and spared no amount of effort to answer our numerous questions. It was largely because of this support that site visits were so enjoyable.

Within AIR, there were others who helped on the project and on this report. Carol Schalla and Ruddie Irizarry made site visits. Jim Shearer assisted in the review of program documents. Carolyn Uyemura, Sharon McVicker, Dorothy Reynolds, Linda Provinzano, and Barbara Mendivil were responsible for typing the report, DRP summaries, and program descriptions. We are also thankful for the guidance provided by staff from the other two studies under this contract.

TABLE OF CONTENTS

Page

PREFACE i

ACKNOWLEDGEMENTS iii

TABLE OF CONTENTS v

LIST OF TABLES ix

SUMMARY xi

CHAPTER I - INTRODUCTION 1

Background 1

Limits of the Study 2

CHAPTER II - BASIS OF THE CRITERIA FOR EXEMPLARY STATUS 5

General Considerations 5

Legislative Implications 7

Criteria for Exemplary Bilingual Education Programs 7

Outcomes 8
Acceptable Evidence 8

Clarity of Instructional and Management Procedures and Availability of Materials 9

Reasonable Start-Up and Recurring Costs 10

Final Considerations 10

CHAPTER III - IDENTIFYING CANDIDATE PROGRAMS FOR THE STUDY 11

Sources of Leads to Candidate Programs 11

Contacting Candidate Programs 11

Acquiring Program Documents 13

Document Receipt and Control Procedures 14

Call/Contact File 14
Card File 14
Document File 15
Alphabetical Notebook 15

CHAPTER IV - APPLYING CRITERIA TO SCREEN CANDIDATE PROGRAMS 17

Document Review and Telephone Follow-Up 17

The Multiple Review System 17

The Rating Sheet 19

Selection of Programs for Site Visiting 20

Comparison Group 21
Measures 21
Sampling (Attrition, Conditions of Exclusion, Size) 21

Levels 22
Statistics 22
Significance 22
Results 23

Evaluation Review by Staff Prior to Site Visits 23

Pre-Site Visit Summaries 24

CHAPTER V - SITE VISITS 27

Purposes 27

Focus 27

Procedures 28

Pre-Site Visit Preparation 28
Site Visit Team Responsibilities and Activities 29
Schedule 29

CHAPTER VI - PROGRAM REVIEW BY THE DISSEMINATION REVIEW PANEL (DRP) . . . 31

Background, Purpose, and Structure of the DRP 31

Preparation of Summary Descriptions for DRP Review 32

Procedures and Results of the DRP Meetings 32

CHAPTER VII - DEVELOPMENT OF DETAILED PROGRAM DESCRIPTIONS 35

Purpose 35

Content and Emphasis 35

Program Overview 35
Program Development 36
Staffing and Management 36
Instruction 36
Parent Involvement and Community Awareness 36
Costs 36
Evaluation 36
Sources for Further Information 36

Review and Revision 37

CHAPTER VIII - DISCUSSION OF RESULTS AND RECOMMENDATIONS 39

Major Deficiencies in Local Program Evaluations 40

Use of Criterion-Referenced Tests 40
Tests of Significance of Gains 40
Misuse of Analysis of Covariance 41
Practice Effect 42
Effects of Revisions of Test Norms 43
Effect of Increased Expenditure Alone 43

Constraints on Implementation and Evaluation 44

Conflicts with Educational Ideals 44
Conflicts with Laws and Regulations 44

Recommendations 45

Guidelines to School Districts on Design for Evaluation . . . 46
Guidelines to Evaluators on Analysis of Data 46

REFERENCES 49

APPENDIX A

Master List of Programs A-1

APPENDIX B

Sample Letters

1. From Terrel Bell, Commissioner of Education B-1
2. From AIR Explaining Project B-3
3. From AIR Reminding Sites to Send Reports (Follow-Up) B-4
4. From AIR Thanking Sites for Cooperating with Project B-5

APPENDIX C

Rating Sheet Used by Program Reviewers C-1

APPENDIX D

Form for Submitting Materials to the Dissemination Review Panel . . D-1

Program Summaries Submitted to the Dissemination Review Panel

Madawaska, Maine D-3
Santa Fe, New Mexico D-15
Philadelphia, Pennsylvania D-27
Alice, Texas D-45
Corpus Christi, Texas D-57
Houston, Texas D-77
Kingsville, Texas D-87

APPENDIX E

Outline for Bilingual Education Program Descriptions E-1

LIST OF TABLES

TABLE — Page

1 Summary of the Results of Identifying Candidate Programs 13

2 Bilingual Education Programs Recommended by AIR and Approved for Dissemination by the DRP 33

SUMMARY

The objectives of this study, conducted by the American Institutes for Research (AIR) under contract to the U.S. Office of Education, were (1) to identify and to recommend as exemplary up to 10 bilingual education programs for submission to the Dissemination Review Panel (DRP) of the Education Division, Department of Health, Education, and Welfare, and (2) to develop detailed descriptions of those programs approved by the DRP for dissemination.

The Office of Planning, Budgeting, and Evaluation (OPBE) and AIR cooperatively defined criteria for screening candidate programs. The criteria were based upon several requisites. To be considered, a candidate program must include English language instruction for children with limited English skills and must teach in the pupils' native language to the extent necessary to allow them to progress effectively through school. Part of the curriculum must also be addressed to the customs and cultural history of the native language group. Furthermore, program participants must show statistically and educationally significant gains in English language skills as well as in the content areas taught in the native language. The program must also have definable and describable instructional and management components. Finally, start-up and continuation costs must be reasonable.

Candidate programs were located from a variety of sources. Initial leads came from the Office of Education and from Title VII files, which were surveyed by the project director at the beginning of the study. Other sources were programs surveyed for past AIR studies, programs reviewed in a previous study by the RMC Research Corporation, regional educational laboratories, state bilingual education officials, and local education agencies. There were 175 programs in the total candidate pool.

A telephone screening was conducted to determine whether programs met minimum requirements to be considered further, and to request that programs send evaluation reports. As a result of the telephone screening, 59 programs were dropped. An additional 20 programs that had been telephoned could not be considered because evaluation reports were not received for them.

The remaining 96 candidates were screened by at least two AIR senior staff members. Information for each program, summarized on a rating form, was the basis for a preliminary judgment about the program. The rating form included sections concerned with the measures, sampling, grade levels, statistics used, and the significance of program effect.

Nineteen programs met the screening criteria sufficiently to warrant further consideration; 11 of these programs were subsequently dropped. Eight programs survived both the initial screening procedure and the following detailed examination and were selected for site visiting.

The purpose of site visiting was to verify evidence of the program's effectiveness as reported in evaluation reports and to observe how the program operated. One program was dropped from the study after the site visit. The remaining seven programs were recommended to the DRP by AIR as exemplary bilingual education programs. Summaries of program operations and evaluation data for the seven were prepared and submitted to the DRP for review. The DRP approved four of the seven for dissemination, as follows:

1. Alice Independent School District
   Bilingual Education Program
   Alice, Texas

2. Aprendemos en Dos Idiomas
   Title VII Bilingual Project
   Corpus Christi, Texas

3. Bilingual Education Program
   Houston, Texas

4. St. John Valley Bilingual
   Education Program
   Madawaska, Maine

Descriptions were written for these four exemplary programs in sufficient detail to provide local educators with a source of ideas to guide them in implementing similar bilingual education practices in their districts. Each program was described as it operated during the year for which the most recent and complete evaluation information was available (in all cases, 1973-74). Some program components were treated historically to illustrate major changes that had occurred since the program began. The descriptions also discussed the context within which each program developed and operated, and the special educational needs of the students.

As a result of the study, AIR made specific recommendations to school districts and program evaluators on ways in which local evaluations could be improved to increase the chances for identifying additional exemplary programs in bilingual education.

CHAPTER I

INTRODUCTION

Background

This document is the final report of one of three studies conducted by the American Institutes for Research (AIR) in the area of bilingual education under contract to the U.S. Office of Education. The study reported here aimed to identify and describe successful bilingual education programs in American schools. It is intended that programs described will serve as models for school systems seeking to improve the quality of the education they offer to children who enter school with little or no ability to communicate in English.

The major tasks of the study were to:

- collaborate with the Office of Education in refining criteria for selecting exemplary bilingual education programs,
- search for programs which met these criteria,
- develop procedures to review and to screen programs according to the criteria,
- visit selected programs to obtain further evaluative and descriptive information,
- submit to the Office of Education and to the Dissemination Review Panel (DRP) of the Education Division, Department of Health, Education, and Welfare, summaries of bilingual programs recommended by AIR, and
- prepare detailed descriptions of the bilingual programs approved for dissemination by the DRP.

The next six chapters of this report correspond to these six major tasks. Chapter II includes a listing, definition, and discussion of criteria used to identify exemplary bilingual programs. Chapter III details the procedures for locating candidate programs and acquiring data and documents needed to screen programs according to the exemplary criteria. Chapter IV describes the review-and-screen procedures for applying criteria to identify the most promising programs for site visits. Chapter V summarizes the purposes, procedures, and results of the site visits. Chapter VI presents procedures for recommending exemplary programs for dissemination and summarizes decisions by the DRP on these recommendations. Chapter VII describes the preparation of detailed descriptions for programs approved for dissemination by the DRP. The final chapter is a discussion of the results of the study and recommendations for improving local program evaluations.

Limits of the Study

A major limitation of this study was its dependence on available, recent evaluation information. Programs that had not conducted evaluations or that did not supply summaries of these data to AIR were not even initially considered. All conclusions reached in this study, then, apply only to those initially identified programs that were contacted by AIR and which supplied program documents to AIR.

A second major limitation was the study's dependence on the review of existing program evaluations. It was not possible to do extensive reanalysis of raw data, nor was it possible to compensate for poorly planned local evaluations. Telephone interviews and, in the case of visited programs, site interviews, had to provide needed clarifications.

The sources from which leads to candidate programs were obtained also limited the scope of this study. The original intent was that nominations be provided by the Office of Education. Sources of possible exemplary programs were mainly lists supplied by the Office of Education and included programs funded under various titles of the Elementary and Secondary Education Act (ESEA) and by the Emergency School Assistance Act (ESAA). As specified in the Request for Proposals, bilingual programs for the handicapped were excluded, regardless of funding source.

A fourth important study limitation was set by the criteria for screening programs. These criteria, established by the Office of Education and refined in collaboration with AIR, intentionally limited the choice of programs to those in which program treatment was consistent with Title VII legislation and definitions, and for which evidence of effectiveness was statistically treated, based on measured achievement and comparison between program participants and some suitable baseline. An example of a disqualification for the former reason would be a remedial reading program for Spanish-surnamed, English-speaking participants who did not need bilingual instruction in order to progress in school. An example of a disqualification for the latter reason would be a bilingual program with evidence of effectiveness that could not justifiably be attributed to the impact of the bilingual program because no baseline or comparison could be provided against which to judge gains made by the participants. In Chapters II and IV in this report, these and other criteria are defined and the results of their application in screening programs are discussed.

Finally, the products of this study--detailed descriptions of bilingual education programs--were limited to programs approved as exemplary by the DRP. The main objective of the DRP review was to scrutinize again the evidence of effectiveness presented for a program and to decide if outcomes were of sufficient importance to make the program worthy of endorsement by the federal government for use by others.

CHAPTER II

BASIS OF THE CRITERIA FOR EXEMPLARY STATUS

General Considerations

Exemplary programs in bilingual education are identified in order that information about them may be disseminated for replication. They are programs that show unusual merit, both in terms of greater than ordinary cognitive and affective gains for their pupils, as well as in the quality of their program elements or components.

Programs judged to be exemplary were recommended by AIR for review by the Dissemination Review Panel (DRP) of the Education Division, Department of Health, Education, and Welfare, which approves the dissemination of information about effective educational practices, products, and projects funded with federal monies. In effect, the DRP issues a federal stamp of approval, labelling as exemplary those educational programs recommended to it that are considered by the DRP to be worthy of replication. Based upon review of program documents, consultation with program staff, and site visits, program descriptions are prepared and submitted to the DRP for review and judgment.

Two major difficulties are encountered in searches for exemplary programs. First, relatively few local program evaluations organize data in a way that permits the results for program pupils to be compared with the results that might be expected otherwise. Second, even when such comparisons are attempted, relatively fewer local program evaluations construct proper design safeguards against alternative interpretations of program outcomes. Consequently, the application of "unusual merit" criteria with respect to cognitive or affective gains results in the identification of only a handful of programs that can be labelled exemplary.

The first task in any search for exemplary programs is to define criteria that can be applied in order to classify those labelled exemplary and those not. These criteria serve as the basis for program evaluation. Since programs are typically described in terms of their components, it is ideally possible to develop evaluative criteria for each of these program elements. Also, these criteria must relate to the objective of replication. Aside from the fact that it is strategically more sound to copy a program with greater than average pupil outcomes rather than one showing ordinary results, replicability potential must also be judged on the basis of other important program features, including reasonable costs, clear instructional processes, and clear management procedures. However, with the exception of general criteria relating to reasonable costs and the clarity with which instructional and management processes can be described, consensus regarding appropriate criteria to evaluate other program processes is rare. Typically, such criteria applied to program processes are those developed and applied during formative evaluation.

Evaluation studies designed to locate exemplary programs favor summative rather than formative evaluation, principally because it is assumed that most developmental problems have been successfully solved as the program matured. Thus most program processes are described but not evaluated; instead, the primary evaluation focus is upon local comparison studies showing that pupils enrolled in the program demonstrate significantly better achievement than would otherwise occur through their participation in a regular or alternative treatment. An exemplary program is one that produces extraordinary outcomes.

Emphasis on this criterion presents a dilemma for the program developer, especially one who attempts replication. The problem is that although program benefits are clear, the relationship between outcomes and program processes is uncertain. When attempts to replicate are unsuccessful, one is unsure how to adapt the replicated program in order to produce the beneficial outcomes obtained at the original program site.

In the absence of known program process-program outcome relationships which permit adaptations to be made in replication attempts, it is necessary to develop ancillary criteria that focus on the generalizability and, hence, the replicability potential of the programs. These criteria involve such factors as the program's duration, its size, its benefits across classes and grades, the appropriateness of measures used to assess its pupils' achievement, its expense, and the clarity with which its instructional and management processes and objectives can be described and understood. Thus, the demonstrated positive effects of an exemplary program upon its pupils must be impressive, and the evidence of effectiveness must be sufficiently generalizable to permit reasonable confidence in the program's potential for replicability.

Legislative Implications

While the above considerations apply in general to all educational programs, the search for exemplary bilingual education programs must also concentrate on the unique characteristics of bilingual education as defined in federal legislation. Bilingual education seeks to change existing more traditional instructional systems in major ways. Its philosophy calls for early instruction in a dominant language other than English. This is an extreme innovation whose effects are transmitted to all parts of any bilingual education program, including needs assessments, teacher preparation, the selection and development of materials, parental and community involvement, and certainly, the formative process evaluation of the program. Ideally, all of these elements would be assessed in a program evaluation.

Federal legislation for bilingual education indicates three key features. First, instruction in the English language is to be provided and English is to be studied. Second, to the extent necessary to allow a child to progress effectively through the educational system, instruction in a child's native language must be provided to children of limited English-speaking ability. Furthermore, instruction in the child's native language is to be provided with appreciation for his cultural heritage, and in all courses or subjects of study which will allow his effective progress through the educational system. Third, a bilingual program must demonstrate effective ways of providing instruction to children of limited English-speaking ability to enable them, while using their native language, to achieve competence in the English language.

Programs that stress only the teaching of English as a second language, or the teaching of a second language to English-speaking children who may voluntarily enroll in the bilingual program do not qualify within the legislative definition of a bilingual program.

Criteria for Exemplary Bilingual Education Programs

As indicated above, a bilingual program is one whose participants are instructed both in English and in their native language to the extent necessary to ensure their effective educational progress. Thus criteria used to identify exemplary bilingual programs must concentrate on the development of English language skills as well as the development of necessary progress-related native language skills. Measures relating to both positive self-concept and to the development of the appreciation of the participants' cultural heritage are important, but are not sufficient by themselves to identify a bilingual program as exemplary.

Outcomes

Evaluations of exemplary bilingual programs must therefore show the following outcomes:

(1) Children of limited English-speaking ability achieve educationally and statistically significant gains in English-language skills in comparison with similar pupils in regular or alternative programs. Comparison results are typically based on program-control group gain contrasts, or show accelerated normative growth over time, or demonstrate improved outcomes for pupils in the programs in contrast to outcomes observed for similar pupils enrolled in the same grade levels prior to the introduction of the program. Measures of these English-language gains will customarily be based upon the administration of standardized tests.

(2) Children of limited English-speaking ability achieve both educationally relevant and statistically significant gains in their native-language skills when instruction in their native language is provided to the extent necessary to ensure their effective educational progress. Evidence of effective progress must be demonstrated and, since common normative measures of content achievement are usually unavailable, acceptable evidence may include teachers' grades or judgments of pupil progress. Grades are rarely reported in summative evaluations.

(3) For children whose dominant language is English and who voluntarily enroll in a bilingual program, reported evidence shows that their development of English-language skills equals that expected if they were instead enrolled in a regular educational program. Acceptable evidence includes test scores, teachers' grades, or teachers' opinions.

Acceptable Evidence

Evidence of bilingual program impact should be based on objective measurements obtained from sizeable pupil samples. Achievement gain measures should be estimated for program participants and for a comparable control group. Well-designed contrasts with pre-program baseline or comparison with appropriate norm reference groups are also acceptable. It is necessary that gains for program participants be significantly greater than gains for the control or comparison group.

Interpretation of the significance of the reported gains depends on customary psychometric and statistical grounds. Measurements should be reliable and valid. Tests should be of appropriate difficulty level for the groups examined. The reporting of achievement in either grade-equivalent or raw-score scales is acceptable; one scale is essentially a linear transformation of the other except at extreme ranges. Average gains for pupils in the comparison groups should be unbiased estimates of the gains for the total population of participants; that is, missing data or the effects of selection should not be great enough to cast doubt on the findings. Confidence in the generalizability and potential for replicability are also greater when results are reported for several classes and grade levels, so that unique teacher or administrator effects can be ruled out.

Statistical significance should be demonstrated so that one may confidently conclude that the results showing superior program effect did not occur by chance; that is, results showing significant program effect, when in fact there is none, should occur no more than five percent of the time. In addition, mean gain differences between program and control groups must be educationally relevant whether reported as grade equivalents or as relative within-group standard deviation units. For example, mean differences of the order of one-half grade equivalent, or one-half standard deviation, are meaningful, as is a large positive shift in mean percentiles between pre- and posttests when program outcomes are compared to those of norm reference groups.

The number of years in which the program has been in existence as well as its expected period of continuation are important program characteristics. A program which has been in existence for several years suggests not only that formative evaluations leading to process improvements have taken place, but also that the critical change from start-up to maintenance management has occurred. Expected longevity suggests confidence in the local district's acceptance of the program and equally important, indicates that the original ongoing program can operate as a resource model for program planners attempting its replication.

Clarity of Instructional and Management Procedures and Availability of Materials

To qualify as an exemplary bilingual program, that is, one whose potential for replicability is strong, it is essential that instructional and management processes be well-defined. Program components and concepts should be presented in a coherent and analyzable fashion structured to assist competent administrators and teachers who attempt its whole replication or who wish to modify existing programs. Instructional materials must be available for dissemination to bilingual educators. The implication here is that copyrights, patents, and other proprietary arrangements should not impede the full dissemination of these materials.

Reasonable Start-Up and Recurring Costs

Program costs are essential items of information for persons planning replication. While it is recognized that variability in total costs and in cost categories exist among programs and locations, it is nevertheless important to provide accurate start-up and recurring budget guidelines for administrative, clerical, teacher and teacher-aide salaries, materials, equipment, facilities, consultants, services, travel, testing, and evaluation. Total per-pupil costs must be within acceptable district limits for replicability, and must be presented in such a way that any add-on costs for bilingual education are clear.

Final Considerations

The decision to recommend a bilingual educational program as exemplary is clearly a summative judgment, arrived at through application of the criteria discussed in the preceding sections of this chapter. Few programs can be expected to show exceptional strengths on all criteria; thus the final decision is based upon a weighing of strengths and weaknesses. Although all programs recommended as exemplary must demonstrate significant achievement gains for their participants, it is important to understand that program components other than merely the program evaluation design are also considered in arriving at a final judgment. These are, as has been discussed, clear instructional and management procedures, availability of materials, and reasonable start-up and recurring costs. These all are essential considerations when judging a program's potential for replicability--which is the crucial reason for identifying exemplary programs in the first place.

CHAPTER III

IDENTIFYING CANDIDATE PROGRAMS FOR THE STUDY

Sources of Leads to Candidate Programs

Leads to candidate programs for consideration in this study were supplied by the Office of Education. It was originally believed that up to 10 exemplary bilingual programs could be identified from this pool. However, additional sources of leads to candidate programs were explored. These additional sources included AIR's extensive program files from a nation-wide search for exemplary reading programs for the National Right to Read Office, the RMC Research Corporation's files from an OE-funded search for exemplary programs in reading and math for disadvantaged children, and nominations obtained in the course of seeking information on current study candidates from regional educational laboratories, state bilingual education officials, and local educational agencies.

The candidate programs identified through these sources included programs funded by Title VII, Title III, Title IV, Title I, and the Emergency School Assistance Act (ESAA). Contacts with Title VII programs were generally limited to those that had been operating at least four years. This period of time was judged to be sufficient for a program to have collected, analyzed, and summarized evaluation data for at least two or three academic years. Also, programs in operation this long were reasoned to be relatively stable. Leads supplied by sources other than Title VII were followed up regardless of how long the programs had been operating, because these nominations had been preselected by officials who were aware of the objectives of the study.

Contacting Candidate Programs

A letter from the Commissioner of Education was mailed to every Chief State School Officer before programs were contacted by AIR. A copy of the Commissioner's letter is contained in Appendix B. The letter informed state education officials about the objectives of the study and indicated that contacts with local program staff would be made in order to obtain information about programs considered for the study.

AIR staff then telephoned project directors or other cognizant local education officials to inquire about each of the candidate programs. A total of 144 programs were called. During these telephone conversations, AIR staff explained the purpose of the study and how the programs being contacted had been identified. The local contacts were briefed on the need for evidence of program effectiveness that would demonstrate success. A minimum requirement for this evidence was that some baseline or comparison be provided so that gains made by the program students could be clearly attributed to the impact of the bilingual program. If program staff indicated that achievement test data and suitable baseline data were available for AIR review, participation in the study was invited.

Initial telephone contacts with non-Title VII programs also included discussion of the program's goals and procedures. Programs were not considered if it was determined that the program focus was outside the scope of this study, e.g., the target group functioned well enough in English that the native language was not required for instruction in other subject areas.

In addition to soliciting the participation of programs in the study, a major emphasis in the telephone conversations was to establish a feeling of good will between program staff and AIR. If programs appeared to be eligible, they were enthusiastically encouraged to participate in the study, but were not pressured if strong resistance or doubts were evident. The very few instances in which such reluctance was encountered were due to previous unhappy experiences with research studies in which information obtained from local officials had been used in ways which were detrimental to the program. AIR staff assured local program officials that in the present study the only programs to be documented and publicized would be those approved as exemplary as a result of AIR screening, OE endorsement, and final scrutiny by the Dissemination Review Panel (DRP) in Washington, D.C. AIR staff pointed out that since this publicity would not be adverse but on the contrary complimentary, it would be advantageous to participate in the study.

Tact was required when the information obtained during the telephone conversations clearly indicated that the program should not be considered further, either because of lack of data, inappropriate focus, or serving a target group that, though bilingual, did not require bilingual instruction. In these cases, the reasons for not pursuing the program's candidacy were explained and agreement was reached with local staff to drop the program from the study.[1] In every case where this was necessary, no hard feelings resulted and local staff were in complete agreement with the decision. A total of 59 programs were dropped from the initial candidate pool on the basis of preliminary telephone contacts or letters requesting not to be considered for the study.

---

[1] Although care was taken to explain requirements of the study and to solicit documents only from appropriate programs, efforts were also made to give the benefit of the doubt and not to discourage any program that might be a potential candidate. As a result some of those that did submit documents also had to be rejected for the above reasons.

Thirty-one programs included in the study were not initially contacted by telephone. These 31 were Title VII programs for which recent evaluation reports were reviewed by the AIR project director at the Office of Education in Washington, D. C. soon after contract award.

Acquiring Program Documents

If answers to telephone queries were affirmative, program staff were asked to forward documents. In several instances, local staff requested AIR to send a confirming letter which described the study, reiterated data requirements, and requested the desired documents. A copy of this letter of confirmation is included in Appendix B.

A follow-up letter was mailed to programs that failed to send the promised documents within 10 days. To encourage programs to send documents even when local files were reduced to single copies, the letter indicated that upon request AIR would xerox and return the reports. A copy of the follow-up letter is contained in Appendix B. A total of 47 follow-up letters were mailed.

Documents were sent by 65 programs; 20 did not send them. Adding the 31 programs for which documents were obtained directly from the Office of Education, the total number of programs for which documents were reviewed was 96. Table 1 summarizes results of identifying candidate programs.

Table 1

Summary of the Results of Identifying Candidate Programs

Initial Candidate Pool		
Programs identified by screening documents at the Office of Education, Washington D.C.	31	
Programs identified on lists supplied by the Office of Education or by leads supplied from other sources	144	
Total Initial Candidate Pool	175	
Drops from Candidate Pool		
Dropped during initial telephone contact or in response to program's written request	59	
Dropped because documents were not received	20	
Total Drops	79	
Remaining Candidate Programs		96

Document Receipt and Control Procedures

The main components of the system for maintaining information files on candidate programs were these:

- Call/Contact File
- Card File
- Document File
- Alphabetical Notebook

Call/Contact File

A folder for each candidate program was maintained in the Call/Contact File. Each folder contained a record of all contacts with the program prior to the receipt of documents, and a sheet indicating documents received. For example, initial telephone conversations with each program's director were summarized in writing on a special form which was filed in the folder. In cases of programs that were dropped on the basis of these conversations the contact form indicates the reason for this action and the name and location of the program contact. If the program asked for a confirming letter before supplying documents, a copy of the correspondence was filed in the program folder. If a follow-up letter reiterating the request for documents was required, a copy of this letter was filed in the program folder. When documents were received, a sheet identifying them was also placed in the folder.

Card File

The main purposes of this file were to identify the number and years of evaluation reports provided by a program and to provide an efficient check-in and check-out system.

When a document was received, a card was typed showing the program title and location and the name and address of the program director or other contact. An identification number was assigned to the program and noted in the upper right-hand corner of the card. Also noted were the academic years for which evaluation reports were sent, and if the program requested that the documents be returned.

Whenever documents were given to an AIR staff reviewer, his or her initials were noted on the document card and were crossed out when the documents were returned. The system was necessary because documents were passed around during the multiple review process, and it was important to be able to tell where they could be found (at any point in time.) For example, when a program's status was being queried, the card was examined to see which reviewer currently had the documents and how many other reviewers had examined them. These individuals could then be consulted regarding the program's standing in the review-and-screen cycle.

Cards were arranged numerically in this file, according to the program identification number.

Document File

The Document File consisted of large manila envelopes, each of which contained all documents for a particular program. On the outside of the envelope was noted the program identification number, program title and location, and the academic years covered by evaluation documents in the envelope. The envelopes were filed numerically by program identification number.

Alphabetical Notebook

As each program card was typed for the Card File, it was xeroxed on a sheet of paper which was then filed in a binder called the Alphabetical Notebook. The sheet for each program was filed alphabetically, first by state, then by site, and finally by program title.

The purpose of this notebook was to enable AIR staff to find the identification number for a program when only the program location was known. For example, if an AIR reviewer needed program documents for the Title VII Bilingual Education Program in Corpus Christi, Texas, he/she first had to know the program identification number because documents were filed numerically by this number. Thus, the reviewer would first check the Alphabetical Notebook under Texas, Corpus Christi, to find the program identification number. Using this number, he/she would look for documents in the Document File. If documents were not there, the reviewer would use the same identification number to locate the check-in/check-out card in the Card File. Documents could then be retrieved from the staff member who had checked them out.

CHAPTER IV

APPLYING CRITERIA TO SCREEN CANDIDATE PROGRAMS

Document Review and Telephone Follow-Up

After program documents had been received and logged in, each program evaluation passed through an intensive review process. Reviewers looked first at the most recent evaluation reports available for each program; in most cases this was for the 1973-74 school year. If a program report for that year was missing and earlier reports looked promising, phone calls were made to obtain more recent reports. If a program had sent only the latest report or interim reports that appeared promising, phone requests were made for earlier reports so that reviewers could examine longitudinal trends in program data. Throughout the entire review, as questions arose that could not be answered by reference to the documents in hand, telephone contacts were made to get specific answers or to request more information by mail.

The Multiple Review System

The screening of program evaluation documents involved a multiple review system. Evaluation reports for a program were given to one of four senior staff members who reviewed them on the basis of the criteria for exemplary status that had been established. Each reviewer evaluated the report independently, using an in-house rating form to record information and judgments under the following categories:

- Comparison method used in program evaluation
- Measures used and content areas measured
- Sampling (attrition, conditions of exclusion, size)
- Levels (grades or ages included in evaluation)
- Statistics used in analyzing data
- Significance (statistical and educational)

(These categories are discussed in detail in a later section of this chapter, and a copy of the rating sheet appears in Appendix C.)

After indicating a yes or no decision on the rating form, reviewers passed the program documents on for another independent judgment. Rating forms were kept by the reviewer who directed the evaluation activities.

Inadequate program evaluations were relatively easy to reject; the better and more comprehensive ones received more intensive scrutiny. When programs appeared promising, or when judgments differed, reviewers met together to compare notes and discuss conflicting judgments.

This procedure was designed to ensure that each program evaluation was selected or rejected on the basis of more than one opinion. However, it soon became apparent that the majority of program evaluations fell short of meeting essential criteria, and that it was inefficient to require independent ratings from four senior staff members when reasons for rejection were so basic that sophisticated analysis was unnecessary. The procedure was therefore modified so that one staff member was given the full-time assignment of reviewing each program evaluation in depth and producing a rating form on which were recorded a yes or no recommendation and, for the guidance of subsequent reviewers, a complete listing of relevant information under each of the six categories, together with references to the page numbers in the reports where that information could be found. Thus the first rating form completed for each program became a resource for all reviewers, even though they continued to make their judgments independently and to use a separate rating form for recording their own notes and opinions.

After this preliminary review, program evaluations were generally passed on to the study's director of evaluation, regardless of whether they had received a yes or no recommendation. The only exceptions were 23 program evaluations with flaws so serious that there was no point in further review. In the majority of these cases, the director of evaluation saw the documents first and made this decision. Shortcomings which were considered serious enough to justify rejection after only one review were the following:

- The evaluation addressed affective aspects only.
- There were no data of any kind, either because the program was too new or because the documents were not evaluations but rather descriptions or proposals.
- Tests were given in only one language.
- No comparisons were possible, either because the program used criterion-referenced tests or teacher ratings only, or used non-standardized tests with no comparison group.
- Evaluation was based on an unacceptably small sample (less than 30), or on a clearly non-target sample (e.g. English-dominant children only).

The project's director of evaluation completed his independent assessment of the documents, using the reference notes of the first reviewer to guide him quickly to the pages where selection procedures were described, measures used were listed, test data and statistical analyses were presented, and results and conclusions were discussed.

If both the first and second reviewers gave a program a negative rating, it was rejected. If both rated a program positively or if their opinions differed, the documents were referred on to another reviewer for his opinion.

The Rating Sheet

A rating sheet was developed for use both as a quick checklist and as an opportunity for brief comment and a preliminary decision (except where a negative decision was clearly appropriate). The headings of the rating sheet are discussed below.

- Comparison group--The specific comparison method used was noted. This could be a control-treatment comparison, a comparison with national norms on a standardized test, or a comparison with typical performances before implementation of the program with growth rates of cohorts prior to the change in method, or a comparison with any other standard--appropriate or inappropriate--used or implied by the evaluators or report writers for the program. If the report contained no specific mention of the method used, but reference was made to standardized tests, to grade equivalents or to percentiles, reviewers would enter "national norms."

- Measures--Under this heading, use of standardized tests such as the Metropolitan Achievement Tests, Stanford Achievement Tests or Cooperative Test Series was noted; also noted were non-standardized tests such as the Inter-American Reading Series (almost the only choice for Spanish), or locally developed tests and criterion-referenced tests. The subjects for which these tests were used were recorded, such as English, Spanish, mathematics, social studies, or science. Affective measures were also noted.

- Sampling (attrition, conditions of exclusion, size)--The generalization of findings, as well as their trustworthiness, are dependent upon the soundness of sampling procedures. Under this heading, reviewers noted the methods used for the selection of treatment groups, and, where appropriate, of the control groups. Reasons for excluding students from the evaluation sample can be as important as those for selection, and for this reason reviewers looked for indications of sample losses and the steps taken to avoid resultant biases in findings.

- Levels--Reviewers recorded here the grade range to which the treatment was applied, e.g., kindergarten through grade 4.

- Statistics--The information noted here covered both the form of the original data (e.g., raw scores, standard scores, percentages or proportions) and the statistical methods and manipulations used for descriptive purposes or for inference.

- Significance--Claims, usually based on the result of F- or t-tests, or of a non-parametric statistical test were noted, along with the reviewer's comments or interpretations. Reviewers recognized that with large samples, trivial differences (in the sense of being of little or no practical importance, e.g., less than a third of a standard deviation) will often reach prescribed levels of statistical significance.

- Decision--The reviewer marked his rating sheet with "yes" if no serious objections to the program were found or with "yes (?)" if he or she had reservations. "No" meant what it said, but "no (?)" meant that the reviewer felt that he or she could perhaps be persuaded to vote for the program if other reviewers found suitably compelling reasons for approving it.

Finally, comments were added to the rating sheet, for example, to obtain telephone clarifications of questions and missing information that were needed before validation review could proceed.

Review sheets with decisions on them were not circulated to other reviewers to avoid influencing their judgments; page references and factual information were provided to speed their review task.

Selection of Programs for Site Visiting

The criteria established for exemplary status (discussed in Chapter II) were applied to screen program evaluations and choose those worth site visiting. In applying the criteria, the following minimal requirements of successful evaluations were considered.

Comparison Group

Some kind of comparison of progress of the bilingual group with that of a non-program group had to be made or implied. Mere demonstration of gains of unspecified size (e.g., significantly higher raw scores at the posttest than at the pretest on a locally developed measure) was insufficient evidence of program effectiveness.

Measures

Although the use of well-known standardized tests of English and mathematics, with acceptable validity and reliability, was preferred, consideration was given to locally developed measures if information on validity and reliability was provided and if progress on this measure could be meaningfully compared with a non-program group. Standardized measures of performance on the languages other than English are simply not available; nevertheless, reviewers examined appropriate measures when these were reported. Measurements of affective attributes (e.g., parent satisfaction or pupils' self-concept) were examined, but were judged to be insufficient in the absence of cognitive achievement data. Criterion-referenced tests were frequently used, but in no case was a program evaluation found which made meaningful comparisons possible.

Sampling (Attrition, Conditions of Exclusion, Size)

There were several important considerations.

- The program group should include at least 40 pupils enrolled in at least two classes. In the end, however, reviewers were unlikely to pass any program with fewer than 100 pupils enrolled in fewer than five separate classes.
- No mention of attrition, or an unusually high rate (say, about 15%), called for investigation. Reviewers were inclined to reject programs with much higher attrition rates, since biases at this rate can exceed a third of a standard deviation.
- Specific exclusion of defined target groups was tentatively regarded as a negative indication, such as confining the reporting of results to students whose English was already fairly good.

Levels

While it was sufficient for a program to operate at a single grade level, in such a case it had to operate in several classes under several teachers. Samples per grade could be relatively small, if several grades were involved.

Statistics

Since processing raw data was outside the scope of this study, the programs reporting lists of students' names with scores were rejected. Data requirements were flexible, however, requiring only that some form of descriptive summaries be given. These could be in raw score units, or in scaled score units. There had to be some measure of central tendency such as mean or median, or a proportion; and there had to be some measure of dispersion, such as a standard deviation, or interquartile range, or some data by which such a measure could be simply derived, for example, by inverting a t-test.

In some statistical summaries, reviewers were able to derive other needed statistics where these were not given. In some evaluations, analyses of variance and/or covariance were reported, not always with justification and sometimes incorrectly. Minor errors were overlooked, but reviewers would not allow sophisticated analysis to cloud issues of poor design or inadequate data. For example, adjustment of means through analysis of variance is common, and suitable for equating small group pretest differences, but examples of adjustments were encountered when control groups differed from experimental groups at pretest by as much as a half standard deviation. If the effect was to make subsequent tests more conservative, it was allowed; otherwise, the program was treated with caution.

Where possible, reviewers looked at the progress of cohorts year by year, and at all grades for a single year. Some progressive losses were expected and found; the data were searched to determine whether these progressive losses were no greater than before introduction of the new program.

Significance

Reviewers looked for the following acceptable possibilities regarding statistical significance.

- The treatment group showed significantly better gains than a control group.
- The treatment group showed significantly better gains than similar students in the same grades prior to the introduction of the program.

- The program students showed significantly greater gains than they achieved prior to entering the program.
- The program students' achievements relative to grade equivalents or percentiles were significantly higher in present grades than they had been in earlier grades without the program.
- The improvement in achievement was large enough to merit attention. Reviewers looked more closely at programs where comparative gains were at least a quarter of a standard deviation or better, although better than a third of a standard deviation was considered minimal for an outcome to be educationally important.

Results

None of the 96 programs reviewed met all the criteria for sound evidence of effectiveness, based on examining their evaluation documents and contacting staff for clarification. Of these 96, only eight programs were finally judged to merit site visiting. The results of program screening are shown in the Master List of Programs in Appendix A, with an indication of why excluded programs were dropped.

The screening process described in this chapter focussed on program evaluation and the quality of evidence supporting program impact on participants. Chapter II also referred to criteria for reasonable program cost and clear and describable program processes. Minimal requirements for meeting these criteria could not be specified as clearly as could requirements for sound evaluation. Furthermore, program documents seldom included cost information, and they presented few, if any, descriptions of instructional and management processes. For these reasons, program processes and costs were not examined closely until site visits were made. The next chapter discusses the site visits to the eight programs selected, and the outcomes.

Evaluation Review by Staff Prior to Site Visits

Program evaluations which passed through the review system described above and received positive ratings (a "yes" or a "yes (?)") from both the first and second reviewers, and a confirmation of this from a third reviewer, were considered to be candidates for site visiting. When several likely programs had been identified, a special evaluation review meeting was held to finalize the decisions as to which should be visited. There were two purposes for such meetings: (1) to allow staff members to raise the sharpest possible criticisms of each program evaluation, so that untenable arguments could be recognized and programs dropped before site visiting, and (2) to point up any questions or needs for additional data which might enhance the arguments in favor of a program, so that site visit teams could collect the information while on site. Questions frequently dealt with the availability of additional baseline data, information about the selection, composition and treatment of comparison groups, and the possibility of obtaining either longitudinal data or the most recent data (in cases where a program had not previously furnished these.) A total of 19 programs had strong enough ratings to warrant consideration at the review meetings.

The review meetings were attended by the four senior staff reviewers, the preliminary reviewer, and the project director. Ideally, all reviewers would have analyzed the program documents, completed rating forms, and noted their questions before these meetings took place. In actual practice, there was not always time for this; however, in all cases at least three senior reviewers had seen the program documents prior to the meetings. The other reviewers and the project director acted as a critical audience for the presentation of the case regarding each program, including both its strong and weak points.

As a result of these meetings, seven programs were dropped from consideration because it was decided that the available data did not justify site visiting and that no better data would be available. An additional four programs that had been tentatively identified for a possible site visit were also dropped after further debate. In the case of the remaining eight programs, it was decided that the evaluation data did justify site visiting and that it appeared likely that additional supporting data could be obtained on site.

Pre-Site Visit Summaries

Once a program had been tentatively selected for site visiting, a brief one-page summary of program information was forwarded to the OE Project Officer, so that he would be aware of basic features of the program and the evidence of its effectiveness. In some cases his review raised additional questions to be pursued on site. In others, he suggested notification of interested state education officials in the event of a visit.

The pre-site-visit summary included the name and location of the program, the year it began, the name of the evaluator and year of the report containing evidence of success. In addition, the following program evaluation features were briefly summarized: comparison method, measures used, sampling, grade levels, aspects addressed by the evaluation, and statistics used. AIR evaluators' comments were added. Finally, the tentative site-visit dates and schools to be visited were indicated.

Pre-site visit summaries were submitted for more programs than were finally visited. Visits to four programs were cancelled because additional information obtained prior to finalizing the visit did not support earlier conclusions regarding their effectiveness. Also, two programs were selected late in the screening process, and there was insufficient time to prepare and submit summaries before the visits.

CHAPTER V

SITE VISITS

Purposes

Site visiting provided the information needed to meet the second major goal of the study--the description of bilingual education programs which were recommended to OE as exemplary programs. There were several reasons for visiting program sites. First, the descriptive and evaluative data included in program documents had to be validated. Second, considerably more detailed information about the management and operation of the program was needed in order to write program summaries for submission to the Dissemination Review Panel (DRP) and to write detailed descriptions of programs that were approved as exemplary by the DRP for dissemination. Finally, visiting the locale, talking with program staff, and observing classroom activities and management techniques provided the perspective necessary to write clear descriptions of program processes.

Focus

AIR site visitors were familiar with the philosophy of and rationale for bilingual education. An extensive literature review, undertaken for a related study under the present contract,[1] had identified variables and issues in bilingual education to be investigated in that study. An indirect outcome was to acquaint AIR site visitors with aspects of bilingual education possibly related to program effectiveness. Discussions of certain key program components appearing repeatedly in the literature did not specify ways in which they might be integrated into a bilingual education program. Site visiting provided AIR staff with insights into how the components interact with each other and with program context as an effective educational process.

To assist site visitors in collecting and organizing data, a detailed outline covering key program components was modeled on outlines used in earlier AIR studies of exemplary educational programs. On the basis of the literature review and suggestions from AIR staff who were knowledgeable about bilingual education, the outline was refined to focus on features considered to be particularly important to bilingual education programs.

---

[1] This related effort was one of three studies that comprised the "Evaluation of ESEA Title VII Bilingual Education Programs" funded under OE Contract No. OEC-0-74-9331.

The outline was used to check the information obtained from documents and to guarantee that adequate information was collected while on site. Because programs vary so widely, however, site visitors were encouraged to allow unique features to emerge during interviews and classroom observations, and to probe in depth those program characteristics that they believed to be particularly significant. Topics covered in the outline were the following:

- Objectives
- Participants
- Locale
- Development
- Management
- Staff
- Preservice/Inservice Training
- Instructional Activities and Curriculum
- Equipment, Materials, and Facilities
- Parental/Community Involvement
- Costs
- Evaluation

Later, to guide writers in preparing program descriptions, this outline was modified further. It is included in Appendix E and described in Chapter VII.

Procedures

Pre-Site-Visit Preparation

After the Project Officer approved AIR's recommendations of programs to be visited, site visit teams were assigned, and preliminary arrangements for the visit were made with project directors, who had been informed of the possibility of a visit by AIR staff.[2] District personnel, usually including the superintendent, were informed of plans for the visit. When requested, background information about OPBE's evaluation of bilingual education and AIR's role in this effort was also sent to superintendents.

Site visitors learned as much about the program as possible from documents that had been sent to AIR. Referring to the outline, they noted areas for which little or no documented information was available. In cases when there was insufficient time before the visit for a thorough preparation, or when limited documentation was sent by the program, questions were formulated from the outline and were sent ahead to the director. This proved to be especially helpful in preparing program staff for AIR's visit and for scheduling interviews with project or district personnel who were knowledgeable about certain areas of program operations.

---

[2] There were two exceptions. In one case, a visit was tentatively scheduled and later cancelled because more recent evaluation reports did not substantiate earlier evidence of the program's effectiveness. In the other instance, further scrutiny of the data revealed inadequate evidence of the program's effectiveness; also, recent events had disrupted program functioning so severely that a site visit would not have been justified.

Team Responsibilities and Activities for Site Visits

In general, each team consisted of two AIR staff members--a writer and a validator--who visited a site for three days. Because it was not possible to send a bilingual person on each visit, AIR staff requested that they be accompanied on site by a bilingual member of the program staff. This individual translated when necessary and explained activities being observed.

The writer collected information necessary to prepare a comprehensive program description. This entailed observing classes in action, reviewing supplementary documents furnished by staff, and interviewing teachers, administrators, and parents. Writers recorded the information in two ways--taking notes and tape recording. They reviewed their notes and tapes each evening and reorganized their questions based on what they had learned that day.

The validator's major responsibility was to confirm the evidence of effectiveness presented in the local program evaluation reports and to analyze additional data that had not been available prior to the visit. The validator used this information about the program's effectiveness to prepare his portion of the DRP program summary. Validators also assisted writers with staff interviews and classroom observations.

Schedule

Visits were made to eight programs. The schedule below shows the site visit dates, location, and title for each.

	Dates	Location and Program Title
1.	February 3-5	La Puente, California--Project BUENO
2.	February 12-14	Houston, Texas--Bilingual Education Program
3.	April 14-18	Alice, Texas--Alice Independent School District Bilingual Education Program
4.	April 14-18	Kingsville, Texas--Kingsville Bilingual Education Program

Dates	Location and Program Title
5. April 16-18	Madawaska, Maine--St. John Valley Bilingual Education Program
6. April 21-23	Santa Fe, New Mexico--ESEA Title VII Bilingual Education Program
7. April 21-25	Philadelphia, Pennsylvania--Let's Be Amigos
8. May 5-7	Corpus Christi, Texas--Aprendemos en Dos Idiomas, Title VII Bilingual Project

During the eight site visits, additional information and documentation were obtained for each program. This new information was reviewed to determine if it substantiated earlier conclusions regarding program effectiveness. In all but one case, it did.[3]

Summaries of the remaining seven programs were written and submitted to the DRP. The nature of these summaries and the results of the DRP review are discussed in the next chapter.

---

[3] One program was dropped after the site visit because the 1973-74 evaluation report, made available during the visit, did not provide evidence needed to support claims that the program substantially improved the English reading and language skills of its participants.

CHAPTER VI

PROGRAM REVIEW BY THE DISSEMINATION REVIEW PANEL (DRP)

Background, Purpose, and Structure of the DRP

The purpose of the DRP was described in a statement prepared by John Evans, Chairman of the Panel, dated 15 November 1974:

> During 1972, the Assistant Secretary of Education and several OE staff members became conscious that the dissemination activities of the various OE programs were proceeding in a largely ad hoc, independent, and uncoordinated manner. . .with the stated or implied recommendation that these materials were "exemplary," "promising," or "effective," and that school systems would be well advised to institute these methods and programs in their own schools.
>
> A review of these disseminations, however, revealed that in most cases there was little or no evidence to back up the claims that the materials were exemplary or indeed effective at all, and that putting an official OE stamp of approval on them and urging schools to adopt them was premature at best and irresponsible at worst.
>
> Accordingly, a Dissemination Review Panel was created to review all programs, projects, models, methods, materials, etc., proposed for dissemination to determine that there was adequate evidence of their effectiveness before dissemination was approved.

The Panel is now a joint body of the Education Division of the Department of Health, Education, and Welfare, with its members drawn from the Office of Education (OE) and the National Institute of Education (NIE). Each agency has 11 representatives on the 22-member Panel. Staff services are provided by the Office of the Assistant Secretary for Education. A quorum of seven members is required for a DRP meeting, and no fewer than three members each from OE and NIE must be present.

DRP members are selected for their ability to analyze and understand evidence of effectiveness supplied by program evaluation data, and also for their general knowledge and experience in education (Evans, op. cit.).

Program submissions are prepared in a special format according to an outline supplied by the DRP. A copy of the outline is contained in Appendix D. (See "Form for Submitting Materials to the Dissemination Review Panel.")

Preparation of Summary Descriptions for DRP Review

A summary of program methods and evaluation results was prepared for each of the seven programs recommended as exemplary by AIR and the Project Officer. The seven summaries are included in Appendix D.

Each summary has two main sections. In the first, key program strategies are described. These are features that appear to be contributing to the outstanding achievement gains of participants as judged from on-site observations, examination of program documents, and interviews with program staff. While DRP judgment is essentially based on the evidence of effectiveness presented in the second half of the summary, the detailed description of key program features must be compelling and detailed enough to lend additional credibility to the data provided and to assure replicability.

The second section of the DRP summary presents evaluation evidence supporting the claim of program effectiveness. The details provided include information on the characteristics of program and comparison groups, identification of instruments used to measure achievement and/or attitudes, and results and interpretation of data analyses. This information is summarized in a manner that highlights the rationale for recommending the program to the DRP.

Draft versions of the DRP summaries were forwarded to the Project Officer for review well in advance of the DRP meetings. The Project Officer reviewed the summaries, recommended revisions that would clarify the information presented, and requested supplementary information on points that might be questioned during the DRP meeting.

Final versions of the program summaries were circulated by the Project Officer to DRP members and Title VII staff in advance of the scheduled review meetings. These summaries were studied in preparation for discussing and judging each program's worthiness of wide-scale dissemination and official endorsement by the Education Division of the Department of Health, Education, and Welfare.

Procedures and Results of the DRP Meetings

The DRP met twice to review AIR's submissions; four program summaries were considered at the first session and three at the second. The presentations were made by the Project Officer who was accompanied by a Title VII representative.

The review procedure involved discussing the information presented in the program summary. Questions from DRP members about specific points were directed to the Project Officer and Title VII representative. Following their discussion, the Panel members voted as to whether the program should be approved for dissemination.

At the first meeting, three of the four submissions were approved. At the second meeting, one of three programs was approved. The programs approved for dissemination by the DRP are shown in Table 2.

Table 2

Bilingual Education Programs Recommended by AIR and Approved for Dissemination by the DRP

Program Title	Location
St. John Valley Bilingual Education Program	Madawaska, Maine
Alice Independent School District Bilingual Education Program	Alice, Texas
Aprendemos en Dos Idiomas, Title VII Bilingual Project	Corpus Christi, Texas
Bilingual Education Program	Houston, Texas

On a percentage basis, 57% of AIR's submissions were approved. This is comparable to the average approval rate of between 50% and 60% for all programs reviewed by the Panel since it was formed. Although a number of bilingual programs have been submitted to the Panel during this time, these four approvals are the first ever awarded to bilingual education programs.

CHAPTER VII

DEVELOPMENT OF DETAILED PROGRAM DESCRIPTIONS

Purpose

A major product of this study is the set of four detailed program descriptions that were prepared for programs approved by the Dissemination Review Panel (DRP). These descriptions were to provide models for project planners, parent advisory groups, teachers, administrators, and others anxious to improve the education of youngsters who have limited English-speaking ability. The program descriptions were written in sufficient detail to permit decisions to be made as to the desirability of attempting locally modified replications of the programs, and indicate sources for further information about each program.

For potential replicators, the DRP seal of approval means that these programs provided evidence of effectiveness contained in their own evaluation reports which was judged to be sufficient to justify their dissemination by the Office of Education as exemplary bilingual education programs.[4]

Content and Emphasis

The Request for Proposals indicated that the detailed program descriptions were to follow an outline expanded from a sample included in the RFP appendix. This outline was initially refined to provide a guide for document review and site visits, as discussed in Chapter V. Subsequently, it was modified on the basis of site visit experience to guide writers to include similar kinds of information in each detailed description.

The outline consists of eight major headings with several subtopics under each heading. Writers tried to follow the outline, and the differences that do occur among descriptions were due primarily to the adequacy of the information available and the unique character of each program.

Below are summarized the content and emphasis of each of the eight major sections of the detailed program descriptions.

---

[4] As discussed elsewhere in this report, programs were not even initially considered for submission to the DRP unless their practices were consistent with Title VII legislation and definitions, and unless evidence of effectiveness met criteria discussed in Chapters II and IV. The label, "exemplary," should be interpreted in the light of these criteria.

Program Overview

In this section the reader is introduced to the program, the impetus for its beginning, program objectives, and general procedures used to achieve the objectives. Also included are staff and student characteristics.

Program Development

This chapter describes in more detail the reasons why the program began, its goals, the program's relationship to the local school system, and aspects of the community that influenced planning of the program. If appropriate, it also indicates people who were instrumental during the planning phase and how the program changed since it began.

Staffing and Management

Included in this section is information about program staff (their qualifications, responsibilities, length of time with the program, recruitment procedures), staff development activities, and program management policies.

Instruction

This chapter describes in detail how bilingual education is implemented in the classroom and includes teaching techniques, grouping patterns, assessment of student achievement, facilities, and materials and equipment used.

Parent Involvement and Community Awareness

In this chapter, the emphasis placed on parent involvement and/or community awareness and examples of how parents are involved with the program are discussed.

Costs

An analysis is given of budget information provided by the program, with special attention paid to start-up expenses and continuation costs. Also given are sources and level of funding. The per-pupil cost for the bilingual program is compared to the per-pupil cost for the regular school program. When possible, budget options are suggested.

Evaluation

This chapter summarizes the evidence of program effectiveness, indicating its impact on student achievement.

Sources for Further Information

In this final section, names are given of program staff who can answer more specific questions about the program; further information may be included on the availability of materials and any references cited in the text are listed.

Review and Revision

After descriptions were drafted and edited by AIR staff, they were mailed to the site for review by program staff. Revisions and changes suggested by them were then incorporated into the descriptions which were submitted to the Office of Education.

CHAPTER VIII

DISCUSSION OF RESULTS AND RECOMMENDATIONS

Programs identified as exemplary in this study were those with demonstrated significant program effect based upon evaluation designs which were sound enough to pass review first by AIR and then by the Dissemination Review Panel (DRP). Criteria for screening the program information have been discussed in earlier chapters. It is important to emphasize that this review was concerned almost totally with the quality of the local summative program evaluations. Per pupil costs and the descriptive clarity of the instructional and management procedures of a program were also examined, but after confidence in the statistical and educational significance of pupil outcomes was established.

This focus on the local program evaluation cannot be viewed as a program evaluation; most program components were not evaluated. In this study, the term exemplary may be strictly applied only to those programs showing significant effect supported by a sound evaluation and approved for dissemination by the DRP. For the remainder of the programs not identified as exemplary, claims of program effectiveness, when these were reported, were not accepted because of inadequate program evaluations. In most cases, one simply could not confidently conclude whether a program was effective or not. As noted earlier, of the 175 programs considered for review as exemplary candidates, 89 reported evaluations so inadequate that no conclusions could be drawn about program effectiveness, 20 supplied no requested evaluation reports, and 59 were dropped after contact because they were too new to have conducted summative evaluations or they did not meet the definition of a bilingual program.

Since few program evaluations met acceptable standards, it is useful to describe some of the more common program evaluation deficiencies in the hope that program developers might be made aware of areas in which their program evaluations can be strengthened. It is probable that many more programs could be identified as exemplary if the results for program participants were properly contrasted with a comparison group, with a relevant norm, or with pupil outcomes prior to the introduction of the program.

Major Deficiencies in Local Program Evaluations

Use of Criterion-Referenced Tests

Many evaluations reported pupil outcomes in which criterion-referenced testing instruments were developed. In no case was one found that demonstrated program effect.

In one somewhat extreme but not atypical example, the program developers had established unreported objectives and constructed evaluative tools "to assess accomplishment for each of the designated subject areas" (no examples or descriptions were provided). Although state-mandated standardized tests were administered to students, they reported no data on these tests. An objective was said to be mastered if 70% or more of the students completed it. The same instruments were used for pretest and for posttest, after which differences in proportions mastering the objectives were tested with a correlated-means t-test. The six pupils in grade 1 improved their mean proportion of correct responses from 0.04 on the pretest for English social studies to 0.09 on the posttest with a t value of 3.21, significant at better than the 1% level. Thus, significance could be claimed although the posttest proportion was far below the mastery level.

Identifying specific objectives, establishing the proportion of pupils who are expected to demonstrate learning of each of the objectives, and checking pupils, objectives, and teaching methods whenever testing shows failure, are profitable uses of criterion-referenced testing. Used appropriately in formative evaluation, this approach can alert program developers to one or more points in a program where attention is needed.

In summative evaluation, criterion-referenced testing is useful when:

- common objectives are set for all schools;
- the scoring system is objective;
- the standards set are tied to agreed-upon typical performances and well-defined conditions instead of being arbitrary.

Tests of Significance of Gains

The most widely-used experimental design is one that presumes to demonstrate program effect through the statistical testing of gain scores (i.e. posttest minus pretest score). The statistical test applied is usually the correlated means t-test, though sometimes a non-parametric test is used. This procedure cannot demonstrate the effects of the program. At best, depending upon the grade level, a grade-equivalent gain of three to five months over a year for a sample of 17 pupils is likely to be significant at the 1% level. This significant gain can be ascribed to maturation, or to practice effect, or to both, and can be independent of the effect of the educational program.

One report contained no less than 204 correlated mean t-tests, mostly of raw scores. Predictably, more than half of these were significant at the 1% level; only 41 were "not significant" in spite of the fact that 60 of the samples contained 14 or fewer students. Every class was tested separately and again as part of the grade level. Nearly half of this evaluation report of over 200 pages was devoted to this type of reporting with most results shown in tables and bar graphs.

The flaw in this approach is not, of course, that there is anything wrong with the statistical procedure itself or even that it does not test the hypothesis proposed; the flaw is that this null hypothesis is not particularly relevant--it states that no increase of learning has occurred over the period. Because of maturation, incidental learning, and the development of test-taking skills, the correct null hypothesis should be that changes in the educational program have brought no change in the rate of increase of learning. This hypothesis is tested when, with appropriate care,

- a control or comparison group is used; or
- comparison is made with the rate of increase in the same group before change in educational conditions; or
- comparison is made with increases in classes previous to educational change; or
- some reasonable basis exists for establishing an expectation of increase in the absence of educational change.

Misuse of Analysis of Covariance

Generalized versions of the t-test are the analysis of variance and the analysis of covariance. When applied to pretest and posttest scores, the same limitations apply as for testing of gains. Analysis of covariance, in particular, is occasionally found misused. This procedure is sometimes used to make adjustments for differences between treatment and comparison groups on pretests. Theoretically, this makes it possible to compare gains of dissimilar treatment groups. When these starting differences are themselves non-significant, such adjustments do little harm. But when the differences are large, this adjustment is improper.

Furthermore, educational measures are interval measures. There is no direct way that the equality of these intervals can be tested, and only the most tenuous way in which rough equivalence can be inferred through the assumption of normal distributed measures. Adjustment through analysis of covariance extrapolates the scale for the lower group upwards, and that for the superior group downwards. There can be no assurance that the two groups are being measured on the same scale, or even for that matter on precisely the same continuum when their score distributions are located at two widely separated regions of the test range.

This should not be seen as a criticism of the analytic process, but of one use to which it is frequently put. In general, though, the indications deriving from analysis of variance seem more useful as starting checks than as arguments for success. They certainly should not be used for major sculpturing of unsuitable data.

Practice Effect

An issue which has received rather scant attention, but which has a potentially strong influence on the interpretation of the outcomes of special education programs is that of the effects of practice on retest scores. In the literature, there are caveats about the increases to be expected from "test sophistication" or from "test interactions" but with no estimates of the size of the effect or of its duration. See, for example, Campbell and Stanley in Handbook of Research on Teaching (N.L. Gage, Ed., 1963, p. 175). Most test users are inclined to dismiss the dangers as not applicable when parallel forms of the same test are used, or as trivial when several months intervene. Both sources of comfort are probably unjustified, but the risks are surely greater when, as in at least one case reviewed, there is deliberate and explicit coaching on home-made parallel forms of the test between testings. In one relevant and suggestive experiment by Verster (1974), the effects of real learning were effectively equalized for four samples who were given an initial test of cognitive abilities, and then given varying numbers of retests at three-month intervals. The first retest for all four groups produced almost identical gains irrespective of the time interval between test and retest. This gain was roughly one-third of a standard deviation; even a lapse of a year had very little effect. The gain for the second retest was also virtually constant for the three groups involved, and of the order of about a quarter of a standard deviation.

If practice effect alone can cause important changes, then part of the amount of reported gains attributed to program effect would have to be discounted. The effect would be largest in the lower grades where there had been little test-taking experience; it would also be largest for less sophisticated students from countries with less emphasis on testing, for example, Portuguese immigrants from the Azores or Spanish-speaking Puerto Ricans.

One program tested its students twice for pretesting, once in Spanish and once with the English form of the parallel test; for the posttest, both English and Spanish versions were again administered. Furthermore, the same students repeated this process in each subsequent grade so that by grade 4 they could have been tested more than 16 times. The comparison group, however, was drawn randomly anew each year with, in all probability, a good deal lower average number of testings.

Effects of Revisions of Test Norms

Over the years, test publishers have sometimes found it necessary to revise their norm tables; this has recently happened to the Stanford Achievement Tests, among others. There appears to be a substantial shift in raw score conversions; the same raw score now reflects a higher grade equivalent, more particularly at the upper grades where the differences can be as much as a full grade or more higher than on the older norms. Whatever the reason, the use of the older norms at first testing or in lower grades, followed by conversions or new norms at retesting or in higher grades, can make the program appear to be successful. When the tests themselves have been revised, the same phenomenon undoubtedly exists, but is then even more difficult to detect or to compensate for.

Effect of Increased Expenditure Alone

Ideally, of course, benefits from bilingual programs should be attributable partly and specifically to the effects of the use of the second language, and increased expenditures should be warranted by this treatment variable. However, the additional funding has been used also for considerable improvement in the facilities and materials for the program group, making it a moot point what combination of program variables caused the positive effect. This, of course, is the central problem in interpretation: the difficulty of specifying which program components significantly predict educational outcomes. To ascribe program effectiveness to just one of its elements is a mistake.

Constraints on Implementation and Evaluation

Conflicts with Educational Ideals

The majority of evaluations reviewed were flawed beyond repair. Applied research is a technical specialty; there is no good reason whatever to expect that teachers and educational administrators are trained and experienced in evaluation. But it is also wrong to assume that problems of evaluation would largely disappear with better training or more use of research specialists. Educational practice and aims are often in direct opposition to the needs of sound summative research. For example, while random allocation of a sample to experimental and control groups is a powerful statistical device, it is virtually impossible in most educational situations; on the contrary, placement in the treatment group is done precisely because there is a need to eliminate a difference between groups. It is difficult for an administrator to consider withholding cases for comparison, since that would produce differences instead of removing them.

Conflicts with Laws and Regulations

Intervention education seeks to minimize differences between performances of groups. Certainly the aim of Title VII appears to be to identify and to assist defined groups of pupils. In at least one case reviewed, this purpose came into conflict with what the court considered to be the objectives of desegregation. Thus, not all failures must be laid at the door of poor design. The following case was not unique, and is an example of problems a program may have to face from federal, state, and court jurisdictions.

Their experimental design was as good as normal educational restraints permit, with use of refined statistical procedures and with clear interpretation of results. They included a control group in their design; checked on the initial comparability of control and bilingual groups; recorded differences in exposure to their treatment; showed the effects of the program on the learning of English, Spanish, and mathematics. They stated their hypotheses before analyzing their results. While evidence for the success of this program was not overwhelming, it was honest and entirely credible.

But then, over a five-year period at least six regional consultants changed the program's guidelines from a planned horizontal expansion to a vertical expansion; control groups were lost; the pattern of bussing was changed, reducing contact between bilingual and English-dominant students. Finally, a desegregation suit caused the closing of one school, a redistribution of

Mexican-American students to predominantly Anglo schools, and a reassignment of teaching staff. The program staff tried to readjust but suffered a drastic cut in staff. They compensated by placing more emphasis on materials development and inservice training, producing 27 specially trained bilingual teachers, all but seven of them paid from local funds--and lost 16 of them to wealthier districts when state bilingual legislation was enacted. The effect of these conflicting and changing decisions must blunt the main thrust of the bilingual program effort.

Recommendations

Finding only seven bilingual program evaluations out of nearly 100 possibilities to recommend to the Dissemination Review Panel was not an unexpected result. The record is no worse than has been found in several examinations of program evaluations carried out in earlier studies conducted for the Office of Education by AIR and others. Practically every study of this type over the years from several research organizations and across a variety of educational programs, including compensatory education, reading programs and now bilingual education, has pointed to poor experimental design, to the lack of planning for evaluation, to inappropriate use of statistical methods, and to a general lack of evidence one way or the other. AIR's experience in this study of finding fewer than 5% of programs which are in receipt of public monies, and have evaluation studies good enough to make closer study worthwhile, is by no means unique. For example, see AIR reports from Hawkridge, Chalupsky, and Roberts (1968); Hawkridge, Campeau, DeWitt, and Trickett (1969); Wargo, Campeau, and Tallmadge (1971); Bowers, Campeau, and Roberts (1974); also see an RMC report by Tallmadge (1974). In these studies, similar standards were applied in screening local program evaluations and a similarly small fraction of candidates survived this close scrutiny.

The findings and conclusions discussed earlier in this chapter suggest several means for improving chances that local evaluators of good projects will be able to develop evaluation reports more likely to satisfy criteria of program effectiveness applied in this study.

It would be a real advance if a substantial reduction could be made in the number of programs now being rejected for lack of evidence, even if this meant an increase in the number disqualified by contrary evidence; this would at least mean an increase in the number to which serious consideration could be given.

Guidelines to School Districts on Design for Evaluation

1. Innovators should be required to incorporate acceptable designs for evaluation in their plans when applying for funds.

2. Many programs will experience considerable difficulty in developing sound evaluations either because they lack the training or because they encounter constraints. They need reassurance that their problems are understood and that even the most difficult circumstances can be made to yield some useful indicators; they may need brief help from a consultant.

3. Guidelines should stress that relatively simple but careful designs can accomplish successfully what the most complex statistical analysis often cannot. Important changes, when their observation is planned, can be demonstrated without resort to a great deal of arithmetic.

4. Some programs are having much difficulty in finding suitable tests in English. For example, monolingual students in fourth or fifth grade who are beginning to learn English could be tested with standardized tests intended for the second grade, but would find the content somewhat beneath them. There is a need for some special tests, or for advice on substitutes.

5. Efforts in test development should perhaps be consolidated to avoid the proliferation of local tests without validity and reliability data or possibilities of comparison. In any case, guidelines or consulting services should be provided to ensure that specially developed tests meet these essential requirements.

6. Many programs could profit from some guidance on the appropriate uses of criterion-referenced testing, and necessity for more uniform criteria when it comes to justification of their funding.

7. Also in connection with designs for testing, most innovators seem to need information on practice effect, and on the records that should be kept if they are to avoid spurious claims of gains. Coaching on tests is another aspect of the same problem, and the limiting conditions under which this is a legitimate device should become more widely known.

Guidelines to Evaluators on Analysis of Data

1. Evaluators, whether from school districts or outside consultants, should outline their evaluations at the beginning of the program. Their needs should be allowed to influence the collection of data.

2. Evaluators should have contact with Title VII, even if only through correspondence. They should be apprised of the kinds of evidence called for,

the weight that such evidence would carry, and the format of their presentations which would simplify the task of readers of their reports.

3. There should be guidelines on acceptable sample sizes, number of classes needed, the attention to be given to attrition and other sample losses, and the conditions under which comparisons can be recognized.

4. Their reports do not need to be overly complicated. Sophisticated analyses can be more nuisance than they are worth if reported in too much detail, or without translation and discursive treatment. Sometimes they are completely superfluous. In particular, analyses of variance and covariance should be treated as tools, not products.

5. Significance testing, particularly of gains in scores between testings, should not have undue importance attached to it. On the other hand, no final report should omit:

- measures of central tendency such as means, medians, or proportions;
- measures of scatter about such points, like standard deviations, quartiles, or ranges;
- some means of translating these dimensions into understandable terms, like norms, stanines, grade equivalents, percentiles, or even units with acquired meaning;
- some benchmark by which the value of the change can be estimated, such as the performance of a previous class, of a comparison group, or of the same group under earlier conditions.

6. Statistical analyses of affective measures or opinion surveys can be developed in addition to those of achievement, but the weight they will carry will usually be small; evaluators should be aware that both the objectivity of the scoring systems and the attention paid to non-responses will be scrutinized before the levels of significance are claimed.

REFERENCES

Bowers, J. E., Campeau, P. L., & Roberts, A. O. H. Identifying, validating, and multi-media packaging of successful reading programs. Final Report. Palo Alto, California: American Institutes for Research, December 1974.

Evans, J. The Office of Education's Dissemination Review Panel: Procedures and criteria. Washington, D. C.: Office of Education, November 1974.

Gage, N. L. (Ed.). Handbook of research on teaching. Chicago: Rand McNally, 1963.

Hawkridge, D. G., Chalupsky, A. B., & Roberts, A. O. H. A study of selected exemplary programs for the education of disadvantaged children. Parts I and II. Palo Alto, California: American Institutes for Research, September 1968.

Hawkridge, D. G., Campeau, P. L., DeWitt, K. M., & Trickett, P. K. A study of further selected exemplary programs for the education of disadvantaged children. Palo Alto, California: American Institutes for Research, June 1969.

Tallmadge, G. K. The development of project information packages for effective approaches in compensatory education. Los Altos, California: RMC Research Corporation, October 1974 (Technical Report No. UR-254).

Verster, M. A. The effects of mining experience and multiple test exposure on performance on the Classification Test Battery. Submitted to the Chamber of Mines of South Africa. Confidential report, C/PERS 220. Johannisburg: CSIR, NIPR, 1974.

Wargo, M. J., Campeau, P. L., & Tallmadge, G. K. Further examination of exemplary programs for educating disadvantaged children. Palo Alto, California: American Institutes for Research, July 1971.

APPENDIX A

MASTER LIST OF PROGRAMS

Appendix A is a master list of all the programs that were contacted as potential candidates for this study. The chart is alphabetized by state, city, and program title, and indicates the language of the target group. The last column shows the result of screening each of these programs, using the following abbreviation codes:

1--The number 1 after a program indicates that as a result of information obtained during the initial telephone contact, the program was dropped without requesting evaluation reports. Many program directors indicated that no data on student achievement could be provided because the program was new or the evaluation cycle was not yet complete. In other cases, initial telephone contacts revealed that the program did not meet the definition of bilingual education necessary to be considered for this project, e.g., the target group already functioned in English and no use of native language was necessary for their academic progress in other subject areas. In a few instances, calls were made as a result of false leads and no special program for non-English-dominant students was operating. Finally, some of the programs were dropped at their own request. The number of programs receiving a 1 was 59.

2--The number 2 after a program indicates that evaluation reports were requested but were not received. The number of programs receiving a 2 was 20.

3--The number 3 indicates that the program was reviewed by AIR staff but that the evaluation methodology was so inadequate that a conclusion about the program's success or failure could not be drawn. Some of the more common shortcomings encountered in reviewing evaluation designs were the following: insufficient or inappropriate comparative data, small numbers of participants and/or control students, unanalyzed data, data reported for one grade level only, inappropriate testing procedures, and failure to collect, in addition to data from language tests, data from tests in other subject areas. In a few instances, program documents supplied little or no information on cognitive achievement of participants. The number of programs receiving a 3 was 89.

REC--This abbreviation means that the program passed the review-and-screen hurdles and was recommended to the Dissemination Review Panel for consideration as an exemplary bilingual education program.

APP--Indicating approval by the DRP, this abbreviation means that the recommendation submitted by AIR and the Office of Education was confirmed after a final review by this group.

MASTER LIST OF PROGRAMS

PROGRAM LOCATION	PROGRAM TITLE	LANGUAGE OF TARGET GROUP	REASON FOR EXCLUSION/ SELECTION *
ALASKA			
Anchorage	Development of Communication Skills with Emphasis on Oral Language and Reading	English	1
ARIZONA			
Ganado	Bilingual Program	Navajo	2
Nogales	Nogales Elementary Bilingual Project	Spanish	2
Phoenix	Phoenix Union High School District Bilingual Program	Spanish	3
Phoenix	Roosevelt District 66 Bilingual Education Program	Spanish	1
Rock Point	Rock Point Bilingual Education Project	Navajo	3
Rough Rock	Bilingual Education Program	Navajo	1
San Carlos	Bilingual Program	Apache	1
Tucson	Bilingual Bicultural Project	Spanish	1
CALIFORNIA			
Auburn	Upper Valley Intercultural Project	Spanish	3
Brentwood	Project Amigos	Spanish	3
Calexico	Calexico Intercultural Design	Spanish	3
Cerritos	Title VII Portuguese Bilingual Program	Portuguese	3
Chula Vista	Project Frontier	Spanish	3
Compton	Title VII Bilingual/ Bicultural Project	Spanish	3
Crescent City	Bilingual Education Program	Hoopa	1
Cucamonga	Title VII Bilingual Bicultural Education	Spanish	3
Culver City	Spanish Immersion Program	English	1
El Monte	Bilingual/Bicultural Education	Spanish	3
Escondido	Bilingual-Bicultural Education	Spanish	3

---

* See page A-1 for explanation of code.

MASTER LIST OF PROGRAMS (continued)

PROGRAM LOCATION	PROGRAM TITLE	LANGUAGE OF TARGET GROUP	REASON FOR EXCLUSION/ SELECTION *
CALIFORNIA (continued)			
Eureka	Northern Indian California Education Project	Karok, Hoopa, Pomo, Yurok, other dialects	1
Fountain Valley	Bilingual Early Childhood Project	Spanish	3
Gilroy	Title VII Bilingual Program	Spanish	3
Gonzales	Gonzales ESL/Bilingual Project	Spanish	1
Happy Camp	Bilingual Education Program	Indian	1
Hayward	Hayward Bilingual-Bicultural Project	Spanish	3
Hoopa	Klamath-Trinity Joint Unified School District Indian Program	Hoopa, Yurok	2
La Puente	Project BUENO	Spanish	3
Los Angeles	Bilingual Schools Program	Spanish	1
Los Angeles	Castelar Bilingual Education Program	Cantonese	1
Los Nietos	Los Nietos Bilingual Bicultural Project	Spanish	1
McKinleyville	American Indian Early Childhood Education-SB 1258	English	1
Montebello	Bilingual Education Program	Spanish	1
Mountain View	Bilingual Education Program	Spanish	1
Orange	Orange Bilingual/ Bicultural Project	Spanish	3
Pasadena	Bilingual Multicultural Program	Spanish	3
Pico Rivera	Bilingual Bicultural Education Para Los Estudiantes de El Rancho	Spanish	1
Pomona	Bilingual/Bicultural Project	Spanish	3
Porterville	Cooperative Effort in Successful Learning Experience--State 1258	English	1
Porterville	Cooperative Effort in Successful Learning Experience—State 2284	Spanish	1
Porterville	Cooperative Effort in Successful Learning Experience--Title VII	Spanish	1

---

* See page A-1 for explanation of code.

MASTER LIST OF PROGRAMS (continued)

PROGRAM LOCATION	PROGRAM TITLE	LANGUAGE OF TARGET GROUP	REASON FOR EXCLUSION/ SELECTION *
CALIFORNIA (continued)			
Riverside	Project UNIDOS	Spanish	3
Rowland Heights	Bilingual/Bicultural Education Program	Spanish	1
Sacramento	Early Childhood Bilingual Education Program	Spanish	3
San Bernardino County	Bilingual-Bicultural Exchange Project (BICEP)	Spanish	3
San Francisco	Chinese Bilingual Pilot Program	Chinese	3
San Francisco	Chinese Bilingual Project	Chinese	3
San Francisco	Project To Advance Cultural Opportunities (PACO)	Spanish	3
San Ysidro	Media Research and Evaluation Center	Spanish	1
Santa Ana	ESAA Bilingual Program	Spanish	1
Santa Barbara	Santa Barbara County Bilingual Project	Spanish	3
Santa Clara	Alum Rock Unified School District Bilingual Education Project	Spanish	3
Santa Clara County	Spanish Dame School Bilingual Project	Spanish	3
Stockton	A Demonstration Bilingual-Bicultural Education Project	Spanish, Cantonese, Filipino	1
Ukiah	Bilingual/Bicultural Project	Spanish	2
Union City	Bilingual/Bicultural Education Program	Spanish	2
Visalia	Allensworth Elementary School Bilingual Program	Spanish	3
COLORADO			
Alamosa	Bilingual-Bicultural Program	Spanish	3
Denver	Bilingual Education Program	Spanish	1
Johnstown	Weld BOCES Bilingual Project	Spanish	3

---

* See page A-1 for explanation of code.

MASTER LIST OF PROGRAMS (continued)

PROGRAM LOCATIONS	PROGRAM TITLE	LANGUAGE OF TARGET GROUP	REASON FOR EXCLUSION/ SELECTION *
COLORADO (continued)			
La Junta	Project Juntos	Spanish	3
CONNECTICUT			
Bridgeport	Primary Bilingual Education Program	Spanish	3
New Haven	Bilingual Program	Spanish	2
DELAWARE			
Wilmington	Bilingual Education Program	Spanish	1
FLORIDA			
Dade County	ESAA-Bilingual Project	Spanish	3
Palm Beach County	ESAA-Bilingual Program	Spanish	1
IDAHO			
Holmdale	Canyon Oahie Bilingual Education Project	Spanish	1
ILLINOIS			
Chicago	Bilingual Education Program-ESEA Title VII	Spanish	3
Chicago	Title III Bilingual Program	Spanish	1
INDIANA			
Gary	Bilingual Early Childhood Education	Spanish	2
KANSAS			
Wichita	Bilingual Education Program	Spanish	1
LOUISIANA			
Lafayette	Lafayette Parish Bilingual Program	French	3
New Iberia	Iberia Parish Bilingual Program	French	3

---

* See page A-1 for explanation of code.

MASTER LIST OF PROGRAMS (continued)

PROGRAM LOCATION	PROGRAM TITLE	LANGUAGE OF TARGET GROUP	REASON FOR EXCLUSION/ SELECTION *
LOUISIANA (continued)			
New Orleans	New Orleans Bilingual Education Project	Spanish	3
Opelousas, St. Landry Parish	Bilingual Education Program	French	1
Ville Platte	ESAA, Title VII, Bilingual/ Bicultural Project	French	3
MAINE			
Madawaska	St. John Valley Bilingual Program	French	APP
MASSACHUSETTS			
Boston	Bilingual Curriculum Project	Spanish	3
Fall River	Fall River Public Schools Bilingual Program	Portuguese	3
Framingham	Bilingual Program of Framingham Public Schools	Predominantly Spanish	3
Lawrence	Lawrence Title VII Project	Spanish	3
New Bedford	New Bedford Title VII Bilingual Program	Portuguese	3
MICHIGAN			
Detroit	Detroit's Innovative Comprehensive Program for Bilingual Students	Spanish	3
MONTANA			
Box Elder	Chippewa-Cree Bilingual Education Project	Chippewa-Cree	1
Crow Agency	Crow Bilingual Education Project	Crow	2
Lodgegrass	Bilingual Program	Crow/English	1
NEW JERSEY			
Lakewood	Bilingual Education in a Consortium	Spanish	2
New Brunswick	Project Better Communication	Spanish	3
Newark	Bilingual Education Program	Spanish	1

---

* See page A-1 for explanation of code.

MASTER LIST OF PROGRAMS (continued)

PROGRAM LOCATION	PROGRAM TITLE	LANGUAGE OF TARGET GROUP	REASON FOR EXCLUSION/ SELECTION *
NEW JERSEY (continued)			
Union City	Title VII Program	Spanish	1
NEW MEXICO			
Albuquerque	Bilingual Education Project, Title I ESEA	Spanish	3
Artesia	Southeastern New Mexico Bilingual Program	Spanish	3
Clovis	Responsive Environment Program for Spanish-American Children	Spanish	3
	Clovis-Portales Bilingual Early Childhood Program	Spanish	
Cuba	Cuba Bilingual Program	Spanish, Navajo	1
Espanola	Espanola Bilingual Education Program	Spanish	3
Gallup-McKinley	Gallup-McKinley County Schools Bilingual Education Program	Navajo, Zuni	3
Hatch Valley	ESAA Bilingual Program	Spanish	1
Las Cruces	Las Cruces Bilingual Education Project	Spanish	3
Las Vegas	ESEA Title VII Bilingual Program	Spanish	1
Mora	Mora Bilingual Program	Spanish	2
Penasco	Bilingual Project	Spanish	1
Pojoaque	Remedial Reading Program	Spanish/English	1
Ramah	Bilingual Education Project	Navajo	2
Santa Fe	Santa Fe Public Schools Title VII Bilingual Project	Spanish	REC
Socorro	Bilingual-Bicultural Project	Spanish	1
Taos	Taos Bilingual-Multicultural Program	Spanish	3
Tucumcari	The Bilingual Program, Title I ESEA	Spanish	3
West Las Vegas	Armijo Bilingual Bicultural Program	Spanish	2

---

* See page A-1 for explanation of code.

MASTER LIST OF PROGRAMS (continued)

PROGRAM LOCATION	PROGRAM TITLE	LANGUAGE OF TARGET GROUP	REASON FOR EXCLUSION/ SELECTION *
NEW YORK			
Brooklyn	Bilingual Program, Title VII School District 13	Spanish	3
Buffalo	Spanish English Developmental Program, Title I ESEA	Spanish	3
NORTH CAROLINA			
Buies Creek	Harnett County Summer Migrant Education Project	Spanish	3
OHIO			
Lorrain	Lorrain City Bilingual Education Program	Spanish	3
OKLAHOMA			
Durant	Choctaw Bilingual Education Program	Choctaw	3
Greasy	Bilingual Education Program	Cherokee	I
Stillwell	Adair County Bilingual Program	Cherokee/ English	3
OREGON			
Woodburn	Woodburn Bilingual Project	Spanish, Russian	3
PENNSYLVANIA			
Lancaster	Educational Opportunity for Bilingual	Spanish	1
Philadelphia	Let's Be Amigos	Spanish	REC
Reading	Title III Modified ESL Bilingual Program	Spanish	3
West Chester	West Chester's Exemplary Program of Bilingual Education	Spanish	3
PUERTO RICO			
Hato Rey	Bilingual Project	Spanish	3

---

* See page A-1 for explanation of code.

MASTER LIST OF PROGRAMS (continued)

PROGRAM LOCATION	PROGRAM TITLE	LANGUAGE OF TARGET GROUP	REASON FOR EXCLUSION/ SELECTION *
RHODE ISLAND			
Providence	Providence's Plan for Bilingual Education	Portuguese	3
SOUTH DAKOTA			
Oglala	Loneman Day School Bilingual Program	Lakota	1
Pine Ridge	Dropout Prevention Program	English	1
Smee	Bicultural Education Program	Lakota	1
West River	Lakota Program	Lakota	1
TEXAS			
Abernathy	Project HABLA	Spanish	3
Abilene	Project ABLE	Spanish	2
Alice	Alice Independent School District Bilingual Education Program	Spanish	APP
Austin	Bilingual Education Program Region XIII Education Service Center	Spanish	3
Beeville	Bilingual Education Program	Spanish	1
Brownsville	Brownsville ISD Bilingual Program	Spanish	3
Colorado City	CC Cable	Spanish	2
Corpus Christi	Aprendemos en Dos Idiomas Title VII Bilingual Project	Spanish	APP
Crystal City	Bilingual Bicultural Education Program	Spanish	3
Dallas	Bilingual and Multicultural Education Program	Spanish	3
Del Rio	Del Rio/San Felipe Bilingual Program	Spanish	3
Eagle Pass	Bilingual Education Project	Spanish	3
Edinburgh	Region One Bilingual Project	Spanish	2
El Paso	Dual Language Program	Spanish	2

---

* See page A-1 for explanation of code.

MASTER LIST OF PROGRAMS (continued)

PROGRAM LOCATION	PROGRAM TITLE	LANGUAGE OF TARGET GROUP	REASON FOR EXCLUSION/ SELECTION *
TEXAS (continued)			
El Paso	El Paso ESAA Bilingual Education Program	Spanish	3
El Paso	Ysleta Independent School District Bilingual Education Program	Spanish	3
Fort Worth	Programa en Dos Lenguas	Spanish	3
Galveston	Early Childhood Bilingual Program	Spanish	3
Harlingen	Bilingual Education Program	Spanish	1
Houston	Bilingual Education Program	Spanish	APP
Kingsville	Kingsville Bilingual Education Program	Spanish	REC
La Joya	Hacia Nuevos Horizontes	Spanish	2
Laredo	Laredo ISD Bilingual Education Program	Spanish	2
Lubbock	Bilingual Education Program	Spanish	1
McAllen	McAllen Bilingual Education Program	Spanish	3
Mercedes	Bilingual Education Program	Spanish	1
Midland	Bilingual Education Project	Spanish	1
Mission	Mission Independent School District Bilingual Program	Spanish	3
Orange Grove	Bilingual Education Program	Spanish	3
Pharr	Pharr-San Juan-Alamo Bilingual Education Program	Spanish	1
Port Isabel	Project We Speak Spanish and English	Spanish	3
Robstown	Bilingual Education Program	Spanish	1
San Angelo	Spanish-English Experience School Program	Spanish	3
San Antonio	Bilingual Bicultural Education Program	Spanish	3
San Antonio	ESAA Bilingual Program	Spanish	3
San Antonio	Harlandale Independent School District Bilingual Program	Spanish	3

---

* See page A-1 for explanation of code.

MASTER LIST OF PROGRAMS (continued)

PROGRAM LOCATION	PROGRAM TITLE	LANGUAGE OF TARGET GROUP	REASON FOR EXCLUSION/ SELECTION *
TEXAS (continued)			
San Antonio	Middle School Program	Spanish	2
	Multi-Media Program	Spanish	2
San Diego	Un Paso Mas Adelante	Spanish	3
San Marcos	Bilingual Education Program	Spanish	3
Weslaco	Project Language	Spanish	3
Zapata	Bilingual Bicultural Project, "Catch UP"	Spanish	3
UTAH			
Blanding	Bilingual Education for Navajo	Navajo	2
Sandy	Jordan School District's Bilingual Education Program	Spanish	1
Salt Lake City	Bilingual Education Program	Spanish	1
WASHINGTON			
Ephrata	Title VII Bilingual Program	Spanish	1
Toppenish	Project BUILD - Bilingual Program	Spanish, Yakima Indian	3
Yakima	Bilingual Education Program	Spanish	1
WISCONSIN			
Milwaukee	Milwaukee Public Schools Bilingual/Bicultural Education Program	Spanish	3
WYOMING			
Arapahoe	Bilingual Program	English	1

---

* See page A-1 for explanation of code.

APPENDIX B

SAMPLE LETTERS

1. From Terrel Bell
   Commissioner of Education
2. From AIR explaining project
3. From AIR reminding sites to
   send reports (follow-up)
4. From AIR thanking sites for
   cooperating with project

DEC 3 0 197

TERREL BELL'S LETTER

DEPARTMENT OF HEALTH, EDUCATION, AND WELFARE
OFFICE OF EDUCATION
WASHINGTON, D.C. 20202

The Office of Education's Office of Planning, Budgeting and Evaluation (OPBE) is conducting an evaluation of the bilingual education program under Title VII of the Elementary and Secondary Education Act (as amended). The evaluation consists of three studies: an "impact" study for Spanish-language bilingual projects which is now in a planning phase, with field data collection to take place during the 1975-76 school year at thirty-five Title VII sites and ten non-Title VII sites; an "exploratory" study for Native American, Pacific and Asian, and European-language bilingual projects, with field data collection to take place during the winter and spring months of the current (1974-75) school year at approximately ten Title VII sites; and an "exemplary" study looking for effective bilingual approaches at approximately ten Title VII or non-Title VII sites, with field visits again to take place during the current school year.

The Office of Education's contractor for this evaluation is American Institutes for Research (AIR) of Palo Alto, California. AIR has prepared a Project Summary for the evaluation, a copy of which is herewith enclosed for your information. The AIR staff, headed by Messrs Malcolm Danoff (Project Director), Antonio de Porcel (Assistant Project Director), Richard Bond (Director of the "exploratory" study) and John Bowers (Director of the "exemplary" study), are presently preparing lists of sites which will be visited and areas of program and policy concerns for which data will be gathered. The OPBE Project Officer, Mr. Edward B. Glassman, is responsible for maintaining contact with the Committee on Evaluation and Information Systems (CEIS) of the Council of Chief State School Officers about this evaluation, which has been included in the annual Data Acquisition Plan.

It is expected that one or more sites in your State will be among the sample groups of projects drawn for the three studies. We would appreciate your cooperation and assistance. I am asking Mr. Glassman of OPBE to provide you with the pertinent list of districts and schools within your State as soon as it is ready. He can be reached at Room 4083, 400 Maryland Avenue, S.W. Washington, D.C 20202, or (telephone) 202-245-7875 for any further information you may need.

I recognize that this is an added request on an already heavy work load in your agency. We do need to do this evaluation in carrying out our responsibilities and hope that it will not be burdensome for your office.

Sincerely,

T. H. Bell
U. S. Commissioner
of Education

Enclosure

cc: CEIS Representative

FORM LETTER FOR PROGRAMS WHO ASKED FOR A WRITTEN REQUEST BEFORE SENDING REPORTS

AIR AMERICAN INSTITUTES FOR RESEARCH

P.O. Box 1113 • Palo Alto, California 94302 • (415) 493-3550 • Cable: AIRESEARCH

(Date)

(Name and Address)

Dear (   ):

This letter is in response to (   ). As part of a larger project in bilingual education being conducted by the American Institutes for Research, we are searching for exemplary bilingual education programs. I would like very much to receive copies of recent program description and evaluation reports that exist for (   ). If you wish, we can copy this information and return your originals.

Following the search for effective programs, AIR will examine their reports in detail and select the most promising for site visiting. Particular attention will be given to the soundness of program evaluation data. A minimum requirement for this evidence of effectiveness is that some baseline or comparison be provided so that gains made by the progam students can be clearly attributed to the impact of the bilingual program.

Dr. Bell, Commissioner of Education, has written to your Chief State School Officer advising him of the project. A copy of the Commissioner's letter is enclosed.

Please call us collect if you have further questions. I appreciate very much your assistance in expediting our request. You may forward the documents directly to our director of evaluation for this project:

A. Oscar H. Roberts
AMERICAN INSTITUTES FOR RESEARCH
P.O. Box 1113
Palo Alto, California 94302

Because our project has an extremely short timeline, we would appreciate it if you could send this information by FIRST CLASS MAIL in the next day or two.

Sincerely,

Project Director

Enclosure

An Equal Opportunity Employer

FOLLOW-UP LETTER

AIR AMERICAN INSTITUTES FOR RESEARCH

P.O. Box 1113 • Palo Alto, California 94302 • (415) 493-3550 • Cable: AIRESEARCH

(Date)

(Name and Address)

Dear ( ):

Several days ago in a telephone conversation with ( ) you kindly agreed to send us evaluation reports for ( ). We have not yet received this information, which is the reason for writing to you now.

If you forwarded this material over a week ago, please call collect to alert us so that we can search our files. Our telephone number is (415) 493-3550. Ask for either Melanie Austin or Sarah Roberts, since they are most familiar with our file of program documents.

If you are sending the material just now, you may forward it directly to our director of evaluation for this project:

A. Oscar H. Roberts
AMERICAN INSTITUTES FOR RESEARCH
P.O. Box 1113
Palo Alto, California 94302

Thanks very much for your attention to this request. Our project is on an extremely short schedule, so you will understand that the sooner we receive the requested information, the more thoroughly we will be able to review your materials. Your reports will reach us much faster if you can mail them FIRST CLASS, hopefully within the next day or two.

Sincerely,

Project Director

An Equal Opportunity Employer

THANK YOU LETTER

AIR AMERICAN INSTITUTES FOR RESEARCH

P.O. Box 1113 • Palo Alto, California 94302 • (415) 493-3550 • Cable: AIRESEARCH

(Date)

(Name and Address)

Dear ( ):

Please accept our thanks for the fine cooperation you extended to AIR in connection with the search for exemplary bilingual education programs conducted by AIR under contract to the Office of Planning, Budgeting, and Evaluation in the Office of Education.

The project is nearing conclusion now and even though your program was not one of those recommended for exemplary status based on evaluation data through 1973-74, you may be interested in more details about the project's purposes, methods, and results and in those programs that AIR recommended as exemplary. This information will be summarized in the final report entitled, "The Identification and Description of Exemplary Bilingual Education Programs."

Again, thank you very much for the help you extended to the AIR project staff.

Sincerely,

An Equal Opportunity Employer

APPENDIX C

RATING SHEET
USED BY
PROGRAM REVIEWERS

RATER______________________

PROGRAM #

COMPARISON GROUP:

MEASURES:

SAMPLING (Attrition, conditions of exclusion, size):

LEVELS:

STATISTICS:

SIGNIFICANCE (Statistical, educational):

Decision: Yes____ No____

APPENDIX D

1. FORM FOR SUBMITTING MATERIALS TO THE DISSEMINATION REVIEW PANEL
2. PROGRAM SUMMARIES SUBMITTED TO THE DISSEMINATION REVIEW PANEL
    - Madawaska, Maine
    - Santa Fe, New Mexico
    - Philadelphia, Pennsylvania
    - Alice, Texas
    - Corpus Christi, Texas
    - Houston, Texas
    - Kingsville, Texas

FORM FOR SUBMITTING MATERIALS
TO THE DISSEMINATION REVIEW PANEL

PROGRAM AREA (e.g., career education, disadvantaged, reading, Title III):

____

I PROJECT TITLE:

II LOCATION:

III SOURCE AND LEVEL OF FUNDING:

IV PROGRAM START DATE:

V BRIEF DESCRIPTION OF PROJECT:

Goals and objectives
Context (community, school, student characteristics)
Program description (grade level(s), years of operation, size, curricula, materials, staffing, facilities, time involved, parental involvement, preservice/inservice training, etc.)
Costs (total, per pupil, initial implementation, ongoing maintenance, etc.)

VI EVIDENCE OF EFFECTIVENESS (in a page or two):

Summarize in appropriate detail the evaluation evidence for the effectiveness of the program or model in question. In order to be acceptable, the evaluation need not be a strict experimental design (i.e., longitudinal measures, random assignment to treatment and control groups, etc.), although this type of evaluation evidence would be the most desirable. However, in order for the Office of Education to officially recommend a particular educational project, technique, or model for wide scale adoption, there must be some kind of high quality, objective, methodologically sound, quantitative assessment which demonstrates that the project in question is effective and superior to other more commonly used approaches or methods. Thus, in order to approve the dissemination of any project, the Panel will require a detailed summary of the relevant evidence including such things as:

-- who conducted the evaluation;

-- sample sizes;

-- improvements or gains in whatever outcome measures were employed;

-- the statistical reliability and educational significance of these improvements;

-- some evidence that the improvements can be attributed to the program and are not just "normal" or "natural" gains (i.e., control group or norm comparisons, or some estimate of what would have occurred in the absence of the program).

PROGRAM AREA:

Bilingual Education (French/English)

PROJECT TITLE:

St. John Valley Bilingual Education Project

LOCATION:

Three cooperating school districts in Northeastern Maine, serving the towns of Madawaska, Keegan, Frenchville, St. Agatha, and Van Buren.

SOURCES AND LEVEL OF FUNDING:

The following figures pertain to the fourth year of program operation.

Title VII funds $175,085

PROGRAM START DATE:

1970

BRIEF DESCRIPTION OF PROJECT:

Goals and objectives. The basic philosophy of the bilingual program is that it exists to supplement the existing program and to assist the community in meeting their linguistic, cultural, social and economic needs for bilingual skills. Designed specifically for this rural area where both French and English are major languages, the program is not viewed primarily as a language program, but contains language only as a medium of the cultural, social, and economic life the students will encounter.

Context. The communities served by the program are small towns in a rural area where the major activity is potato farming. Madawaska, the largest, also has a paper mill, which processes wood pulp pumped across the river (also the U.S.-Canadian border) from its larger neighbor, Edmundston, Canada. Before the boundaries were drawn this was the area settled by the French-speaking Acadians. There is much commerce, social interaction, and intermarriage between the Canadian and American communities, and the Acadian heritage remains a vital influence in the area. Edmundston is primarily French-speaking, and many people in the American towns along the other side of the river also still speak French at home and use both languages in their daily life. Some children come to school speaking French better than English, and poorly prepared for school in either language.

Program description.

Grade levels, years of operation, size -- During its fourth year of operation, 1973-74, the program served 768 children in grades K-4.

Staffing -- The program staff required for this number of participants consists of the following: one full-time Project Director, one full-time Curriculum Coordinator, one part-time French Consultant, one full-time Evaluator, 36 full-time teachers, and 22 full-time teacher aides. All but a few are bilingual.

Curricula, material, time involved -- Designed as a supplement to the regular curriculum, the bilingual program added a French element which was completely new, and also added some new elements to the English curriculum, primarily a greater emphasis on speaking and vocabulary building. Thus, children now have certain basic skills and concepts of the regular curriculum presented to them through the medium of two languages, plus a special series of topics in French which relate to their cultural background.

Since the program operates in three districts whose curricula, staffing patterns, and physical arrangements were independent, there are differences in the approach among the various schools. Grouping patterns may differ and English curriculum materials and activities may differ. However, all bilingual program classrooms share two basic and essential characteristics: They use a sequence of instructional objectives developed and refined by the teachers for each grade level over the years of program operation. They use standard procedures for evaluating achievement of these objectives. The evaluator visits all classrooms frequently to guide and consult with the teachers in assessing student achievement; this helps to maintain uniformity across schools. The feedback from teachers and their involvement and responsibility for the curriculum and objectives, are important characteristics of the program.

In addition to the uniform objectives and evaluation procedures, the program has developed a special series of French curriculum materials; although not mandatory, they appear to be used in all classrooms. Teachers get a folder for each objective which contains the section of the curriculum guide for that objective. They generally build on this, adding their own materials and activities. The French curriculum is built on the fact that the students live in a stable, rural area, with limits on experiential situations. The commercial materials used in the regular curriculum seemed to have shortcomings in teaching certain things because they referred to situations totally unfamiliar to the students. The French curriculum is used partly to remedy this. If the students' test results show that they are weak in a certain basic area of math, for example, that topic will be stressed the next year in the French component. For this reason, some objectives were revised annually during the first four years, on the basis of test results. The content areas addressed through French instruction include mathematics, music, art, social studies, and language skills. About two-thirds of the instruction takes place in English, one-third in French.

The French curriculum materials, based on the students' "cultural context", were planned during the first year of the program to follow a logical sequence over the elementary years. Expanding outward from the student himself, the topics are as follows:

Grades K-1 -- the child, his family and friends
Grade 2 -- the town and local area
Grade 3 -- the State of Maine and the Canadian Province of New Brunswick across the river
Grade 4 -- New England, Canada, and Louisiana (all homes of the Acadians)
Grade 5 -- the continent of North America

Materials for Grade 6, presently being developed, will reflect a human development emphasis, focusing on the various parts of the world where French is spoken.

Facilities -- The program can operate in a regular classroom without any special modification. The arrangements actually used vary from open classrooms created by knocking out old walls to self-contained classrooms with desks in rows.

Parental involvement -- Parental involvement in the program has grown over the years, as parents, some initially fearful that their children would not receive a strong background in English, observed that the English curriculum was not being weakened and liked the introduction of activities in French which brought important features of local life into the children's schooling.

The parents' organization has played an active role in getting the community involved in the program. They carried out community surveys of parent attitudes, and they solicited and contributed articles by parents for the project bulletin issued eight times a year. Parents have also helped with hosting visitors to the program and have served on an occasional basis as volunteers in the classrooms.

Preservice/inservice training--Training has been conducted for teachers and aides each year since the program began. Workshops have been offered by visiting consultants from universities in the U.S. and Canada, by members of the Title VII program staff, by staff members from other bilingual programs, and by publishers of materials used in the program. These sessions, although optional and held outside of regular school hours, have had near 100 percent attendance by both teachers and aides since the program began. The inservice program has also included summer institutes, courses offered through the University of Maine at Fort Kent, visits to other systems and to other schools within the program, and attendance at national conferences. The aides have participated in training sessions together with the teachers, and through a special study program at the university they earn credit toward a degree in education. Attendance at the University courses is also voluntary, and an average of 20 to 25 teachers and aides have participated in each course.

Subjects of the training sessions have included classroom management courses where the aides and teachers define their roles working together in classroom situations. Other topics have been Acadian history, preparing objectives, new experiments and trends in bilingual education, use of specific materials, evaluation procedures, developing English language arts activities, and information brought back by staff members who participated in conferences or visits.

Cost--During Year 1 the Title VII program received funds totaling $150,000; 400 students were served. Major start-up costs included $26,000 for instructional materials and equipment; once classrooms are equipped it is assumed that only a minor expense for new materials will be required for some years afterward. For 1973-74 the cost was $175,085 to serve 768 children. The $228 per-pupil cost can be compared with $375 per pupil in Year 1. Regular district per-pupil costs for the three districts were $495, $465, and $567 in 1970-71 and, respectively, $817, $637, and $727 in 1973-74. One reason for the decrease in per-pupil cost of the bilingual program is that although the number of children served has increased, the program administrative staff has not. Also, anticipating the end of Title VII funding at the end of Year 5, the local districts have begun a move to continue the program by gradually picking up the salaries of the aides. The Title VII expense for aides has reduced only somewhat, but many more aides are now employed. The major costs to the districts for continuing the program without outside funding would be the salaries of aides and administrative staff, and the price of instructional materials.

EVIDENCE OF EFFECTIVENESS:

Evaluation

An unusually high standard of evaluation and reporting was set from the first year of the program in 1970-71 by Heuristics Inc. and Mr. Gilman Hebert jointly, and maintained since then by Gilman Hebert. Also unusual was the fact that "the services of the evaluators . . . was engaged before the start of classroom instruction; therefore evaluation of the project was both formative and summative in nature." While we do not always agree with these evaluators on the relevance of a few of their demonstrations (e.g., we set little store on the raw score gains from pre- to posttests) their statistical analyses are always straightforward, appropriate, and fully discussed. The most recent report (1973-1974) was produced by Gilman Hebert, with the technical assistance of Heuristics.

Reported are the results of standardized tests, and the attainment of product and process objectives, from kindergarten to grade 4. A particularly useful feature in the most recent report is a table showing longitudinal data for 1970-1974 for Total Mathematics and Total Reading for grades K to 4; and the graphical presentations of distributions of stanines for each standardized subtest for each grade represent an additional refinement.

Each report has, for each section, a chapter on "Commendations and Recommendations" which reflect shortcomings noted (as well as successes) with suggested action for the next year.

An unusual feature of this program is that it covers three distinct school districts with whole-hearted collaboration, and a single evaluation report covering each district separately, as well as their combined results.

Sampling

Different procedures were adopted for selection of program students. In Frenchville and St. Agatha (S.A.D. #33) all students in kindergarten of Dr. Levesque School were included in the Title VII program. In Madawaska itself, only those students entering kindergarten who were volunteered by their parents were included; and in the third district of Van Buren (S.A.D. #24) entering students were randomly assigned to classes, and classes were then directed to program or non-program. Language dominance was not considered in allocation, so that French-dominant, English-dominant and bilingual students were retained in successive grades, with only minor changes either way at the request of parents.

Six schools, two from each district were involved; ultimately 60 percent of all students were participants; there were 768 students, 36 teachers and 22 teacher aides in this last year of reporting. The table below is taken from the report. The last row in the table shows ranges, in percentages, of incomplete data for all tests.

Table 1

Enrollment in Bilingual Program in 1973-1974, by School and by Grade

School District	K	Grade 1	Grade 2	Grade 3	Grade 4	All Grades
Frenchville and St. Agatha (SAD #33)	19	56	55	55	60	245
Madawaska	43	65	75	74	43	300
Van Buren (SAD #24)	46	46	50	38	43	223
Total	108	167	180	167	146	768
Attrition %	7	11-14	14-18	7-10	11-13	10-13

It should be noted that these procedures do not assume random allocation to program or non-program. Comparisons are made with national norms, and no other comparison or control group was identified in the report, but the superintendent's offices of the three school districts were able to give us additional data, which allowed us to make some comparisons with the non-program students. This will be dealt with later.

Measures

The new (1970) versions of the Metropolitan Achievement Tests have been used since the start of the program, for all grades except K, with all subtest scores being given. These were given with Form F as pre-test in September, with the alternate Form G as a posttest in May, i.e., with seven months between the two. Levels used were Primary I for grade 1, Primary II for grade 2, Elementary for grades 3 and 4. Specifically, progress in Reading, Language and Mathematics were reported.

In grade 1 only, the Common Concepts Foreign Language forms I and II Test in French was given as pre- and posttest, with their own local tests for other grades, and with achievement of specific objectives also being recorded for each student at every grade level.

The Durrell Listening Reading Test is also given in grades 2 and 4 in January. The SRA tests are applied throughout the Madawaska school district and Frenchville.

Comparison Procedures

The evaluators prepared to make three kinds of comparisons:

- gains from pre- to post-test,
- with national norms,
- with preset criteria of objectives.

Understandably they avoided the possibility of comparisons between schools or school districts and reported their data for total samples by grade. Although other standardized tests were applied in the schools, only the Metropolitan was routinely applied in all districts. The main comparisons were therefore restricted to this instrument, and to specially constructed local tests and objectives. The latter of course could have no wider implications than local.

For kindergarten, the MAT Primer Level yielded standard scores and stanines. Distributions on the latter were graphically plotted against a background of a normal distribution in addition to a table.

For the higher grades, a table for each grade showed, for each subtest, the size of the sample, means and standard deviations for pre- and post-tests, using both standard scores and grade equivalents. In addition for each subtest, the stanine distributions were shown graphically superimposed upon a normal distribution.

Oral ability in French was tested at grade 1 with the Common Concepts Foreign Language Test, French Version, with gains from Form I as pretest, to Form II as posttest. Then the sample was divided into English-dominant and all others. Using raw scores, the evaluators made comparisons over the four years of existence of the program.

As mentioned earlier a single table showed the trends and cohort performances on the MAT for Total Reading and for Total Mathematics, across grades for the duration of the four years of the project. The results were given in all three measures, Standard Scores, Stanines, and Grade equivalents.

For their specified criteria in English and French, tables showing percentages of students attaining all the objectives were given for each grade.

Lastly, we did receive some data on which we could compare performances of program and non-program students, for SRA tests mainly.

Results

The most revealing table is probably this extract from their longitudinal study.

Table 2

Results in Stanines and Grade Equivalents on the Metropolitan Achievement Tests, for Total Reading and Total Mathematics by Grade Level and by Academic Year

Grade	Subject	1970-1971		1971-1972		1972-1973		1973-1974	
		Stanine	GE	Stanine	GE	Stanine	GE	Stanine	GE
K	Reading	-	-	6	-	7	-	7	-
	Mathematics	-	-	6	-	6	-	7	-
1.8	Reading	5	1.8	5	1.8	6	2.0	5	1.9
	Mathematics	5	1.9	5	1.8	5	1.8	5	1.9
2.8	Reading	-	-	5	2.6	5	2.8	6	3.1
	Mathematics	-	-	4	2.5	6	3.0	7	3.3
3.8	Reading	-	-	-	-	4	3.1	5	3.5
	Mathematics	-	-	-	-	5	3.8	6	4.2
4.8	Reading	-	-	-	-	-	-	5	4.3
	Mathematics	-	-	-	-	-	-	5	4.7

Comparing achievement with placement, there seems to be a slight loss in grade 4; but for the rest, in 1973-74 average grade equivalents are close to, or even above placement.

Following the transit diagonally, of the 1970-1971 grade 1 group, it will be seen that they started at par; lost two months in Reading and three in Math in grade 2 (1971-1972); lost an additional five months in Reading but caught up in Mathematics in grade 3 (1972-1973); and finally lost no more ground in Reading and only a trifle in Mathematics in grade 4 (1973-1974).

The next cohort did better. In grade 1, (1971-1972) they were placed as they had achieved in both subjects; in grade 2 they were if anything a shade ahead in Mathematics; in grade 3 they had lost a trifle in Reading, but were four months ahead on Mathematics.

The third cohort (grade 1 in 1972-1973), starting again at par, were ahead on both subjects by the next year (grade 2 in 1973-1974).

There is therefore a suggestion that the program may be refining its techniques and improving its achievements over time. Also, at entry the average ability for the program is if anything a little above the average ability of the population (stanines of 6 to 7); thereafter the average performance in Reading is a shade below this promise (stanines generally at five, with one four and two sixes), with Mathematics a bit better (stanines of four, several fives, a six, and one seven). When it is remembered that these children are for the most part French-dominant, this seems a good performance, and improving over time. In addition, they are of course taking an extra subject - French.

The evaluators for this program followed the common practice of testing the significance of gains from pretest to posttest; they did so using the correct t statistic for correlated means and as usual these differences in every case proved to be highly significant. As is our usual practice, we discount the value of such demonstrations.

On the other hand a most useful device has been their graphical presentation of distribution of stanine scores against the background of the normal distribution. This was done for every subtest for each grade. Almost uniformly these show slight upward shifts, with a suggestion of peaking, or narrowing of the scatter, and occasionally with a hint of negative skewing as if the ceiling of the tests was making its presence felt.

For every grade, in every subject (including Art, Music and Social Science) objectives had been defined and for each the percentage of students attaining this was given. There were about 140 of these objectives, spread fairly evenly over the five grades. For ninety of these, the percentage of students attaining the objective was 80 percent or higher. For kindergarten, one value was below 70; only 38 percent satisfied the criterion on phonetics in French. The evaluators were less satisfied with the program's effects in grades 3 and 4 where about a quarter of the percentages were below 50, most of them for Music and Social Science.

Lastly, at our request during our site visit, we were given additional data for Madawaska, which allowed us to make some comparisons for the Title VII group with the rest of all non-program students.

In two schools in Madawaska, CTMM values of IQ were obtained at the grade 3 level in November 1973 and in November 1974. The table below shows the results.

Table 3

IQ Means and Standard Deviations (approx.) in Grade 3 for Title VII and Regular Classes for 1973 and 1974

Year		Title VII	Regular Classes	Total Group
1973	Mean*	111	104	107
	SD*(±)	13	15	14
	N	69	58	127
1974	Mean*	110	102	106
	SD*(±)	17	17	17
	N	76	42	118

*These values were derived from quartiles and medians, and are therefore approximate.

These differences are statistically significant at about the 1 percent level, but of course could reflect, at least in part, the effect of two years in the program for the Title VII students. However it is important to note that, especially in 1974, the standard deviations are if anything larger than usual; had there been explicit selection, we would expect all SD's to be reduced to well below 15 to about 9. The fact that the SD for the total group is about the same as that for each of the subgroups shows that there must be considerable overlap between program and non-program distributions. More than half of all students are in the program.

Also for Madawaska, we obtained the results of SRA tests applied in October 1974 to both Title VII and regular classes, in grades 2, 3 and 4. The results are shown below.

Table 4

SRA Grade Equivalents for Title VII and Regular Classes, Grades 2, 3 and 4

		Title VII		Regular Classes	
Grade Placement	Subject	Mean*	SD*	Mean*	SD*
2.9	Reading	2.4	1.2	1.9	0.6
	Language Arts	2.6	1.2	2.0	0.4
	Math	2.6	0.7	2.1	0.7
3.9	Reading	3.7	1.0	3.0	1.0
	Language Arts	3.9	1.3	3.1	0.7
	Math	3.4	1.2	2.9	0.7
4.9	Reading	4.7	1.3	4.0	1.0
	Language Arts	5.0	1.6	4.3	1.2
	Math	4.6	0.9	4.1	1.0

*These values have been derived from quartiles and medians, and are approximate.

Here again we can see that the Title VII students are sometimes a trifle behind grade placement in their average achievement, though there are no cumulative losses. They are far ahead of the Regular Classes in every case. It is not possible to say for sure whether this is due to initial biases as a result of the self-selection (parents' choice), or whether it reflects a superiority accumulated as a result of two or more years in the Title VII program. However there is one phenomenon that should be taken note of: While the standard deviation of grade equivalents for the regular classes is close to our expectations (usually about 2/3 of a year in grade 2, rising to over one year at grade 4), the values for the Title VII are higher. This is contrary to what we would find if selection had taken place, but quite in keeping with the hypothesis that the treatment has expanded the ranges upwards. There is certainly no evidence of handicap for these students, and they are learning French.

Comparisons between Title VII and Non-Title VII pupils was also made possible in the other two districts by additional data. For these we have grade equivalents, but no standard deviations.

School district Van Buren used the Metropolitan series of tests throughout the schools. Besides Title VII, they also have the Follow-through program, but were able to supply the following data for these subgroupings and for two grades.

Table 5

Mean Grade Equivalents on MAT for Grades 1 and 2 with Subgroupings, for May 1974

		Title VII				Non-Title VII			
Grade Placement	Subject	Follow-through	(N)	NFT	(N)	Follow-through	(N)	NFT	(N)
1.8	Reading	1.6	(22)	2.0	(21)	1.3	(42)	1.5	(18)
	Mathematics	1.5	(22)	2.0	(21)	1.2	(42)	1.3	(18)
2.8	Reading	2.7	(21)	3.5	(12)	1.9	(46)	2.2	(37)
	Mathematics	2.9	(21)	3.3	(12)	1.9	(46)	2.0	(37)

The superiority of the Title VII students is patent; the lower of their subgroups (Follow-through) have higher performances than the better of the Non-Title VII subgroups (Non-follow-through).

For Frenchville we have SRA test results for grades 2, 3 and 4. In grade 4, by special request from a teacher with a group of 20 low ability students, her class was included with 22 other students in the Title VII program; this of course depressed the mean performances given below.

Table 6

SRA Mean G.E.'s for Title VII and Non-Title VII Groups

Grade Placement	Subject	Title VII	(N)	Non-Title VII	(N)
2.9	Reading	2.4	(70)	1.8	(47)
	Mathematics	2.4	(70)	2.1	(49)
3.9	Reading	3.6	(73)	3.2	(59)
	Mathematics	3.5	(73)	3.1	(59)
4.9	Reading	3.7	(42)	4.6	(88)
	Mathematics	3.6	(42)	4.5	(87)

It should be remembered that grade 4 included a special class of low-ability students. For the other grades the Title VII students are ahead.

Comments

The impression created, after the site visit, was of thoroughness of both the educational and the evaluation procedures. It was a pleasure to find no trace of political or ulterior motivation, but on the contrary a clear determination to avoid handicap for their students in English, without losing their historical and predominantly French origins and culture. Every classroom had evidence of considerable emphasis on conceptual development in both languages, and we frequently heard the class switch from French to English to reinforce some non-language theme, as for example in social studies or mathematics.

The data and statistics are more solid than startling, and the more believable for that. They are quite in keeping with what we saw and with the competent professional atmosphere we found.

PROGRAM AREA:

Bilingual Education (Spanish/English)

PROJECT TITLE:

ESEA Title VII Bilingual Education Program

LOCATION:

Santa Fe, New Mexico

SOURCES AND LEVEL OF FUNDING:

The following figures pertain to the fourth year of program operations and do not include the local funds normally allocated by the Santa Fe Public Schools for the education of the children who are served by the bilingual program.

Title VII Funds	$ 72,281
State Funds	58,615
Total Year 4	$130,896

PROGRAM START DATE:

1970

BRIEF DESCRIPTION OF PROJECT:

Goals and objectives. The primary goal of the program is to develop children with a good self-concept who can communicate and learn in both languages, English and Spanish. Part of the bilingual philosophy is to maintain the language and culture of the Spanish heritage in the Southwest; at the same time, giving children a good education in English is viewed as essential to the role of the schools.

Context. The Santa Fe area still reflects strongly the traditions of its earliest inhabitants, the Indians and the Spaniards. Within the city itself, there is a considerable second-language influence of Spanish. Accurate measurement of the predominance of Spanish has been difficult, although the Santa Fe Public Schools have made various attempts. The actual proportion of students whose primary or home language is Spanish probably falls somewhere between the 63% based on Spanish surnames and the 34% reported by a survey of parents. The three Title VII schools average 80% Spanish surname; all three are also Title I schools.

Program description.

Grade levels, years of operation, size--During its fourth year of operation, 1973-74, the program served 419 students in 19 classrooms. Of these, 8 were first-grade classes, 5 were second grade, 3 were third grade, and 3 were fourth grade. Twelve of the classes, one at each grade level in each of the three original Title VII schools, were included in the evaluation. The other 7 classes were part of a horizontal expansion into new schools.

Staffing--The program staff reauired for this number of participants consists of the following: 1 full-time project director, 1 full-time materials specialist, 1 full-time secretary, 1 part-time evaluator, 19 full-time teachers, and 19 full-time aides. All instructional personnel are bilingual.

Curricula, material, time involved--The children served by the Santa Fe Bilingual Program are predominently Spanish-surnamed, but very rarely Spanish monolinguals. Generally, they come to school speaking both Spanish and English rather poorly, with Spanish the primary language for at least a third of them. The bilingual program adds to the regular English program a Spanish instructional component which complements and reinforces the instruction in all content areas. Thus students receive a bilingual presentation of all the topics of study in the normal curriculum. For example, a teacher might introduce a certain math concept in Spanish in the morning, then reinforce the lesson in English that afternoon. In language arts, the regular state-mandated English curriculum of the Santa Fe Public Schools is followed. In addition, Spanish language instruction is provided, end parallels the English instruction so that objectives are at roughly the same stage in both. A child should generally be reading at about the same level in both English and Spanish, for example.

The program uses Spanish curriculum materials from the Spanish Curriculum Development Center (SCDC) in Miami, and is participating in the "regionalizing" of these materials by suggesting revisions which make them more appropriate for the Southwest. Four strands of the SCDC curriculum are used: science-math, language arts, Spanish, and fine arts. The materials include behavioral objectives and criterion-referenced tests. In addition, the teachers have developed special social studies objectives which are locally meaningful. Thus, although the Spanish curriculum parallels the regular English curriculum, it also adds topics which are especially relevant to the program children and their background. For instance, the Spanish words and historical figures which gave names to Santa Fe streets are included in the second-grade study of communities.

A bilingual materials center has been established to serve the teachers, providing them with SCDC and other Spanish-language materials, and with special supplementary materials in English. The materials specialist consults with teachers about materials they need, and is also in the process of bringing together the many materials developed by individual teachers, which will be selected, catalogued, reproduced and made available to all program teachers.

Classrooms are divided into "interest centers" where children work on the various subjects in the curriculum. These centers are stocked with materials in both languages, and are equipped with various audiovisual and independent study aids, such as language masters, listening center, typewriter, and filmstrip viewer. For language arts/reading, pupils are grouped by ability; in all other subject areas groupings are flexible according to each child's needs and interests. Generally, the teacher works with children in one of the room's centers; the aide works with a second group, and the rest work independently, alone or in small groups. The frequent shifting of groups and subjects complements the shifting of the language of instruction, as the teacher conducts a lesson with one group in Spanish, and they progress to an independent activity in either language, and then perhaps to an English lesson with the aide in another center.

Facilities--The program can operate in a regular classroom; the only modification required is the rearrangement of furniture to create the different centers.

Parental involvement--Over the years, a very active Parent Advisory Committee has taken the responsibility of organizing parents for a variety of activities in support of the program. Parents have visited the legislature to convey their approval of the program, have accompanied children on field trips into the community or the nearby pueblos, and have volunteered their help for special occasions in the classroom, displaying crafts and customs, and preparing food. On one such occasion, parents brought various types of traditional Lenten foods during Lent, and they, their children, and the teachers gathered in the gymnasium to share the meal.

The Parent Advisory Committee also coordinated special public events to acquaint the community with the program. One of these was a series of meetings in which a panel of experts discussed bilingual education, members of the Bilingual Program staff talked about Santa Fe's program, and demonstrations were presented by children and teachers. Another event was a citywide bilingual program commemorating Cinco de Mayo, in which children from the program presented traditional dances and celebrations on a Sunday afternoon.

Preservice/inservice training--Teachers have had an important role in developing and refining the program. They have written behavioral objectives each year as successive grades were added, and have revised these in subsequent years. Beginning with the second year of operation, when grades one and two were included in the program, inservice training for teachers and aides was offered through universities in the area. The district contracted with the University of New Mexico in 1971 and with New Mexico Highlands University in 1972-75 for the provision of inservice training courses for which academic credit would be given. The courses were carried on throughout the school year, and included such topics as development of behavioral objectives, assessment of entry levels, specific teaching techniques in the skills and concepts areas, using Spanish in teaching, materials development in Spanish, Spanish culture in the Southwest, and a seminar and practicum in bilingual education focusing on components of the Santa Fe program.

Cost--During Year 1 the program expended $58,598 and served 77 children in three first-grade classrooms; during Year 4 total program cost was $130,896 for 419 children, 19 classrooms, grades 1-4. These 19 classrooms include one class at each grade level in each of the three "pilot" schools where the program was first installed under Title VII; these are the classes included in the evaluation. In addition, parent requests prompted the horizontal expansion of the program into 7 more first- and second-grade classrooms in 5 other schools. Here, as in the pilot schools, the program adds a grade each year. Start-up costs for each new classroom were primarily for materials and equipment and the aide's salary. Continuing costs have included the salaries of aides and of the program administrative staff, inservice training, evaluation, and general support services. Because the expenses of the program administrative staff have remained relatively constant as the number of children served has increased, and because much of the equipment lasts for more than 4 years, the program's per-pupil cost has decreased from $761 in Year 1 to $312 in 1973-74. This is in addition to the regular district per-pupil cost of $803 for 1973-74.

In 1973-74, funds were allocated as follows: Title VII funds covered the third- and fourth-grade classrooms in the pilot schools. Grades 1 and 2 in both pilot and expansion schools were covered by state bilingual education funds. During the coming year, the Santa Fe Public Schools plan to pick up part of the program cost as its expansion continues, an estimated $18,000 to pay for aides.

EVIDENCE OF EFFECTIVENESS:

Evaluation. This was done by Dr. Charles F. Leyba of California State University, Los Angeles, and submitted for the academic years 1972-73 and 1973-74. The earliest reports for 1970-71 and 1971-72 were done by Dr. Mari-Luci Ulibarri. The most recent report is the most comprehensive in terms of data and analyses. It is a substantial effort covering the year's performances by grades 1-4, for bilingual groups and their control groups at each level. The bulk of the report consists of the data for each and every pupil on every test given, with statistical summaries.

Of these summaries, results of criterion-referenced testing are fully reported for both English and Spanish, mainly in the form of proportions or percentages, with some interesting comment by the evaluator. He also reported on the results of standardized tests using more refined statistics; these cover only pages 7-12 but, supplemented by a little additional analysis by AIR, are the most revealing of the effect of the program.

It should be noted that the earliest report by Dr. Mari-Luci Ulibarri threw light both on deficiencies in experimental design, and on shortcomings of the program; in both areas her report appears to have had the attention of the school district organizers with good results in the next year. In particular, her comments on the use of covariance analysis for adjustments of differences in control and experimental groups, and on the need for random selection of assignment, are models. While random assignment to control and bilingual groups was not done, the sample used for control was randomly selected from all pupils not placed in the bilingual group; this yielded some improvement for the comparisons which followed, which were in the main, for each grade level, of performances by control and bilingual groups on each subtest of the Metropolitan Achievement Tests separately, for Total Reading and Total Mathematics, and for attainment of objectives in English and Spanish.

Sampling. Three schools were involved. In each, for each of grades 1-4 (New Mexico does not have kindergarten), all students in the bilingual (i.e., treatment) groups were used, while random samples of about 24 (stratified by school) were drawn for each of these grades from all students not placed in bilingual groups.

It should be noted that division into bilingual and non-bilingual groups was not random. Parents could elect to have their children placed in either group, and to have them transferred at later stages. This process is likely to have caused biases though we made one attempt to check on this, as will be seen later. One hypothesis we were not able to test was that perhaps at this stage parents whose children were already fairly fluent in English wanted their children to learn more Spanish, while those whose children had difficulty with English preferred to concentrate on it. While this hypothesis did not seem to hold in the earliest stages of the program, a recent phenomena, superiority in English from the beginning of schooling, is difficult to account for without it. In the control group for each of the grades between 21 and 25 students were initially placed, with attrition losses generally of four or less; grade 3 lost 5 out of 25. Each of the bilingual groups started with 65 or more, with a maximum attrition, in grade 1, of 18%; in grade 4 only 7 out of 68 did not complete all tests. The table below gives fuller detail.

TABLE 1

Samples Used in 1973-1974

Grade	Bilingual		Control	
	Initial	Used	Initial	Used
1	65	53	24	24
2	66	58	21	18
3	72	62	25	20
4	68	61	23	22
TOTALS	271	234	93	84

Losses were the result of incomplete data on either pretest or posttest, or both.

Measures. The Metropolitan Achievement Test forms F and H, appropriate levels, were used for pre- and posttests.* Criterion-referenced measures for Spanish and English are also reported in the form of "numbers of items correct on each test." All subtests for the MAT are reported separately, so that performance in mathematics can also be evaluated.

In addition, scores on the Otis-Lennon Mental Ability Test are available. This test was applied to each pupil at entry to school.

Comparison procedures. Form F of the MAT was given as pretest to each of the subgroups during the fall (September) and form H was applied as posttest in spring of the following year (May). Thus, comparisons possible for each grade separately are raw score means with standard deviations only:

- Bilingual group with control group on pretest scores
- Bilingual group with control group on posttest scores
- Bilingual group with control group on gains
- For each group, gain from pretest to posttest

The criterion-referenced tests also yield comparisons of bilingual with control groups for each grade and on both pre- and posttests. These results are given as percentages for each school and for both languages.

---

*With the lower overall performance it was found to be better to extend the use of each test level by half a year at the posttest in order to avoid floor effect.

Results. It is understandable that evaluators often seek to compensate for differences in starting levels of treatment and control groups by comparing gains rather than posttest scores only. The objections to the comparison of gains are even more cogent than those given by the first evaluator of this program to adjustments through analysis of covariance; we will not go into this here, but while we have quoted the gains in tables 2 and 3 below, we have added the simpler comparisons. These tables are derived from the 1973-1974 Evaluation report.

TABLE 2

Total English Reading Means for 1973-74 on Metropolitan Achievement Tests

	Bilingual			Control			Differences		
Grade	Pre-test	Post-test	Gain	Pre-test	Post-test	Gain	Pre-test	Post-test	Gain
1	15.34	27.72	12.38	12.35	24.21	11.96	2.99*	3.51**	0.42
2	46.31	64.19	17.88	43.44	58.00	14.56	2.87	6.19	3.32
3	45.77	64.49	18.72	51.10	64.16	13.05	-5.33	0.33	5.67
4	47.97	60.97	13.00	41.18	53.68	12.50	6.79	7.29	0.50

TABLE 3

Total Mathematics Means for 1973-74 on Metropolitan Achievement Tests

	Bilingual			Control			Differences		
Grade	Pre-test	Post-test	Gain	Pre-test	Post-test	Gain	Pre-test	Post-test	Gain
1	12.90	26.23	13.32	10.38	24.92	14.54	2.52*	1.31	-1.22
2	34.83	51.66	16.83	30.67	39.89	9.22	4.16	11.77**	7.61
3	54.91	74.61	19.70	52.70	66.05	13.33	2.21	8.56*	6.37
4	51.17	67.56	16.39	44.77	65.59	20.82	6.40	1.97	-4.43

*Indicates two-tailed significance at between 1% and 5% levels.

**Indicates higher orders of significance.

Standard deviations and sample sizes were given for every group mean above, as well as for differences of gains (last column), so that while the evaluator quoted significances for these gain differences only, we were able (since we discount the value of such demonstrations) to calculate the significance of differences of posttests. Only the Total Math posttests differed significantly ($P < 0.1\%$); but of the 16 test differences, only five were significant at the 5% level or better, and only one was negative. Thus the bilingual groups are at the least holding their own or better, while of course getting instruction in Spanish in addition. The superiority of the bilingual group over the control at the pretests from grade 2 onwards could be the effect of treatment received up to that time.

Attention is drawn to the Total Reading performances of the two grade 1 groups where for both pre- and posttest the bilingual group had significantly higher means. For the pretest, at least, this is noteworthy since there is no earlier treatment to account for the superiority (New Mexico has no kindergarten). We were provided with Otis-Lennon mental ages for all children in the two groups by Mr. Romero, and found these summary statistics:

	Bilingual group	Control group
Means (Months)	73.56	72.96
Standard Deviations	10.68	8.42
Sample Size	52	24

(Note: One pupil in the bilingual group had no test result here)

The difference of 0.6 months in favor of the bilingual group is in the right direction, but quite insignificant and surely not large enough to account for the differences in Total Reading, which are of the order of a half standard deviation. This seems to point to some bias in the bilingual group; indeed it seems reasonable for parents whose children are already fairly fluent in English to seek the addition of Spanish instruction, and for those whose children need it, to prefer additional attention to their English. If this hypothesis could be sustained, it is in any case a recent development; in the first two years of the program, the control group was considerably superior to the bilingual group.

The results of the criterion-referenced tests will not be given in detail here. In the evaluation report, data is given in the form of percentage passing a test. Results are given for each of 12 to 17 tests, both as pretests and posttests, for bilingual and control groups for each grade, in each of the three schools, in each of the two languages. Results varied from 0 percent to 100 percent in almost every subdivision, making summary difficult. However, in Table 4 we have the averages of percentages passing the 12 to 17 tests in each subgroup. It is not possible to condense the table further.

TABLE 4

Averages of Total Percent Correct in Criterion Tests, Pre and Post, in English and Spanish Versions for Each Grade in Each of the Schools Participating in 1973-74, for Control and Bilingual Groups Separately (Derived from 1973-74 Evaluation Report)

	GRADES							
	One		Two		Three		Four	
English Version	Pre	Post	Pre	Post	Pre	Post	Pre	Post
Alvord								
Control Group	55	76	30	61	31	47	46	58
Bilingual Group	65	81	43	87	32	90	74	81
Larragoite								
Control Group	71	89	38	43	20	35	55	63
Bilingual Group	81	87	48	75	30	90	57	73
Agua Fria								
Control Group	73	90	48	69	24	36	49	56
Bilingual Group	71	84	52	85	31	50	49	63
Spanish Version								
Alvord								
Control Group	37	54	18	41	10	13	34	26
Bilingual Group	42	81	36	89	16	71	63	80
Larragoite								
Control Group	41	52	26	28	7	24	19	22
Bilingual Group	46	69	35	65	22	93	30	57
Agua Fria								
Control Group	31	47	30	37	4	17	13	15
Bilingual Group	40	82	44	85	23	41	25	38

In the first column in Table 4 are the averages of the percentages of students passing the 12 criterion pretests in each group in each school in grade 1. It should be noted that once again the bilingual group has better results than the control group in every case but one. As this occurs before treatment, it would seem that the bilingual group is more select than the control group. This bias may be present in later grades, although differences between groups appear to increase at the posttest in English. At least the bilingual group appears to suffer no handicap in English. In the Spanish versions the bilingual groups generally do much better than the control groups.

In Table 5 below we give the posttest grade equivalents for grades 2, 3, and 4 for each of the two groups.

TABLE 5

Grade Equivalents of Means of Posttest (Metropolitan Achievement Tests)

	Total English Reading		Total Mathematics	
Grade Placement	Bilingual	Control	Bilingual	Control
*1.8	(61)	(46)	(50)	(41)
2.8	2.33	2.25	2.66	2.03
3.8	2.96	2.95	3.14	2.74
4.8	3.92	3.52	3.95	3.87

* Grade equivalents are not available for grade 1; percentile ranks are given instead. Table derived from 1973-1974 Evaluation Report.

We can draw these conclusions from the last three tables and the reports on the criterion-referenced tests:

- Both groups are losing ground each year--a common finding.
- The bilingual group is ahead of the control group at the posttest for every grade for both Total Reading and Total Mathematics.
- The bilingual group increases its lead in Spanish every year.
- If anything, the superiority of the bilingual group is even more pronounced for mathematics than it is for reading.

Comments. We should have liked to have included longitudinal studies since the program has been in existence since 1970, so that the first grade 1 pupils are now in grade 4, but it would have involved us in further manipulation of data with not enough time.

Upon visiting the classes (and we visited every class in all three schools), the satisfactory comparisons made occasioned no surprise. All three schools set high standards of cleanliness, orderliness, and of energetic and purposeful performance by all principals and teachers. All classes from which the control samples were drawn were examples of traditional teaching at its best, with never more than about 25 pupils per teacher, and with a good deal of evidence of both training and effort on the part of the teachers; but the bilingual groups had in addition an aide in every classroom and all the modern ancillaries of language masters, projectors and overhead projectors, tape recorders, etc. It would have been a surprising and interesting phenomenon if these things made no difference. And on the other hand, knowing that the bilingual group had Spanish as an additional course of study, it would have been suspicious if the differences were larger than they were; that this did not happen is almost certainly due to the careful experimental design, which excluded many sources of bias.

NOTE: This program is recommended for two major reasons. First, the evaluation outcomes show that most objectives are being met or surpassed. Second, the evaluation design is exemplary and replicable. See pages 8 through 18.

PROGRAM AREA:

Bilingual Education (Spanish/English)

PROJECT TITLE:

Let's Be Amigos

LOCATION:

Philadelphia, Pennsylvania

SOURCE AND LEVEL OF FUNDING:

The following figures are for the fifth year of operation.

Title VII Funds	$695,066
Other Federal Funds	245,000
Local Funds	705,000
Total Year 5 Funds	$1,645,066

PROGRAM START DATE:

1969

BRIEF DESCRIPTION OF PROJECT:

Goals and Objectives. The Let's Be Amigos Program is designed to

- increase students' first and second language competencies,
- provide for conceptual development in the students' mother tongue,
- increase students' self-esteem,
- locate and/or develop curriculum materials relevant to the needs of Philadelphia's Puerto Rican students,
- identify and develop the talents of Spanish speaking community members, and
- systematically monitor and improve upon previous program activities.

Context. Program students attend four high schools, two junior high schools, and five elementary schools located in the inner area of a port of entry city. Of the 1,830 Spanish language dominant students (mainly Puerto Rican) and the 1,000 English language dominant students served by the program in 1974, 100% came from families with an average annual income under $4,000. Program schools are characterized by over-crowding, high absenteeism, and a 40% turnover in student body each year.

Program Description

The Let's Be Amigos Program consists of three educational components: Model A, Model B, and Arriba. These components are distinguished by the populations they serve, by their staffing patterns for instructional and support personnel, and by the time they devote to each language. On the other hand, all three components have similar curriculum philosophies, parent-community involvement, inservice-preservice training, and program monitoring. In the remainder of this program description, the aspects which distinguish the three components are described for Model A, Model B, and Arriba in turn. Then similarities among the three components are summarized under the headings Curriculum and Materials, Inservice-Preservice Training, Parent-Community Involvement, and Program Monitoring.

MODEL A

Grade levels, years of operation, size -- Model A began in 1969 in the newly constructed Potter-Thomas demonstration school for bilingual education. During the 1973-74 school year, all 760 Spanish-dominant students and 640 English-dominant students in grades pre-K through 4 at the school were served. (One of the kindergarten classes is an all-day kindergarten for more advanced students.)

Staffing (instructional and support) -- In Model A a bilingual, Spanish-dominant teacher and a bilingual, English-dominant teacher work together to provide instruction for 70 students who are working at the same grade level. The Spanish-dominant teacher teaches only in Spanish and is the homeroom teacher for the Spanish-dominant students. The English-dominant teacher teaches only in English and is the homeroom teacher for the English-dominant students. Students usually stay with their homeroom class when receiving instruction in language arts, reading, social studies, and mathematics. If certain students seem ready to read in a second language before others in their class, they may join more advanced classes during the language arts period. Students from the English-dominant and the Spanish-dominant homerooms meet together weekly for such manipulative activities as science experiments, art projects, dancing, and singing.

Non-teaching support personnel at the Potter-Thomas demonstration school include two reading specialists, one mathematics specialist, a bilingual librarian, a media specialist, and a program coordinator. The subject-matter specialists plan demonstration lessons, order new materials, help place students in the proper classes, and supervise program teachers. The program coordinator also serves as liaison between the bilingual program staff and the school principal, who is not bilingual.

Time involved -- In general, pre-kindergarten and kindergarten students in Model A receive a half-hour of instruction daily in their second language. Students in the all-day kindergarten class, grade 1, and grade 2 receive 1 hour of instruction per day in their second language.

Spanish-dominant students in grade 3 receive 2-1/2 hours of their instruction in each language, while English-dominant third grade students receive 1-1/4 hours of instruction in Spanish and 3-3/4 hours of instruction in English.

According to their linguistic ability, students may receive special help from two "English as a Second Language" teachers, one "Spanish as a Second Language" teacher, and two Spanish reading teachers. One full-time aide and one part-time aide are also available to help each team of teachers provide more individualized instruction.

MODEL B

Grade levels, years of operation, size -- Model B began in 1970 as a replication of Model A in other schools. During the 1973-74 school year, Model B served 150 Spanish-dominant and 80 English-dominant students in grades 1 through 6 at Ludlow Elementary School, as well as 100 Spanish-dominant and 50 English-dominant students in grades 1 through 3 at Miller Elementary School. Students, parents, or school officials must request that a student participate in the Model B program, as there is only one Model B class per grade level at each school.

Staffing (instructional and support) -- Teachers in the Model B program are bilingual Spanish-dominant, or bilingual English-dominant. There are no aides, no Spanish reading teachers, and no bilingual program, non-teaching support personnel at the Model B schools. However, Model B teachers may go to Potter-Thomas School to consult with support personnel as needed.

Time involved -- Each day Model B teachers are assisted for 45 minutes by an English as a Second Language teacher and for another 45 minutes by a Spanish as a Second Language teacher. These two periods are used to provide instruction and study in English as a Second Language and Spanish as a Second Language, respectively. During the rest of the day, Model B teachers try to devote equal time to instruction in both languages in social studies, mathematics, sciences, and supplementary activities.

ARRIBA

Grade levels, years of operation, size -- Arriba began in 1969 and was designed to serve newly immigrated students who speak virtually no English, as well as English speakers interested in Puerto Rican history and culture. During the 1973-74 school year, Arriba served 1,050 students in grades 4 through 12 at three elementary schools, two junior high schools, and four senior high schools. School counselors, parents, and teachers recommended these students for placement in the Arriba program.

Staffing -- At the elementary level, one bilingual teacher is responsible for from 20 to 30 students. If the bilingual teacher is English dominant, she may provide 1-1/2 hours of English as a Second Language instruction for her students. If the teacher is Spanish-dominant, she may team with an English-dominant teacher for English as a Second Language instruction. The Arriba teachers do not receive any further instructional support or assistance.

Time involved -- At the elementary level, 15 to 20 Arriba students are placed in ungraded classrooms with two bilingual teachers. Each day, the English-dominant member of the bilingual teaching team provides 1-1/2 hours of English as a Second Language instruction for these students. All other classroom activities are in Spanish.

In the junior high and high school Arriba programs, students also receive 1-1/2 hours of English instruction a day. Students also may choose to study social studies, Spanish language arts, commercial subjects, mathematics, science, and music classes conducted entirely in Spanish. Course content for these classes is planned to parallel the regular school program, and Arriba students join non-Spanish-dominant students in physical education, music, art, health, and black history classes.

Curriculum and Materials

A full-time curriculum coordinator, four full-time teachers on special assignment, and 22 teachers working part time as curriculum developers consult with administrators and teachers to determine what materials should be adopted, adapted, or developed to serve Model A, Model B, and Arriba students.

Commercial materials adopted by the program include Laidlaw and Lippincott basal reading series, Schneider's science series, Finocchiaro and Lado's English as a Second Language materials, Addison-Wesley's Mathematics in Spanish, and numerous audiovisual aids.

Adopted materials include translation into Spanish of locally developed social studies units, the development of an audio-lingual approach to beginning oral Spanish, and phonics instruction to accompany the Laidlaw reading series.

Materials developed to fit local program specifications include a series of short stories written to reinforce the learning of consonants, and flash cards to teach vocabulary and usage.

In general, language learning materials used by the Let's Be Amigos program stress the use of the audio-lingual method (frequent repetition, pattern response drills, memorization of dialogs, and translation only when necessary). Materials also tend to be action-oriented and include suggestions for songs, dances, and art projects. Audiovisual materials including film strips, flash cards, and cassette players are also recommended by the curriculum development staff.

Inservice-Preservice Training

Preservice and inservice training for the Let's Be Amigos program is conducted by program supervisors, professors from Temple University, outside consultants, the program director, and program teachers. Twice a year, two series of staff development meetings are offered to all bilingual program staff members. To encourage attendance, participants are paid. A series might be devoted to such areas as English as a Second Language, Spanish as a Second Language, science, mathematics, or classroom management. Occasional meetings for all district personnel are scheduled to acquaint school system staff with the bilingual program and with the history and culture of the Puerto Rican. A typical meeting might feature a speech by the President of the University of Puerto Rico, videotapes of bilingual classroom sessions, or a locally developed film on Puerto Rican history and culture.

Each teacher in the bilingual program is observed on an average of once a week by her bilingual program supervisor, four times a year by her principal, and intermittantly by research staff members and school-based reading and mathematics specialists. After each of these visits the observer and the teacher meet to discuss methods, materials, and classroom management techniques.

Native Spanish speakers who have completed 60 semester hours of college, who have taught in other countries, or who have participated actively in community affairs may apply to attend eight-week summer institutes co-sponsored by the bilingual program and Temple University. The Pennsylvania State Department of Education issues intern teaching certificates to participants who successfully complete the institute. These participants then attend summer and evening classes to fulfill requirements for a regular teaching credential.

Parent-Community Involvement

Each week the Coordinator of Special Projects and a teacher on special assignment meet with a different group of from four to six parents. Meetings are held at one of the parent's homes. At these meetings, the Coordinator of Special Projects or the teacher on special assignment explains the bilingual program's goals, answers parents' questions regarding bilingual education, and encourages parents to attend monthly Advisory Committee meetings. Parents are allowed to bring their children to Advisory Committee meetings, and local entertainment follows most of these meetings.

The Coordinator of Special Projects and the teacher on special assignment also serve as liaisons and interpreters when parents wish to speak with principals. In one instance, the result of such a meeting was to obtain a bilingual parent to lead the home-school association. To keep the community informed about the program, the Coordinator of Special Projects produces a weekly radio show entitled "Let's Be Amigos." The program format features student talent, talks with parents and local citizens, and is geared to community interests. Although it does not focus on program publicity, the exposure given to the Coordinator enhances her effectiveness as a program leader.

Program Monitoring

Each summer the head of the research division discusses results of the previous year's evaluation with bilingual project staff members. He then helps the staff members write revised, measurable program objectives for the coming year. The head of the research department and his two assistants then supervise program implementation on a rotating basis. They visit the bilingual program schools on an average of twice a week to supervise testing, to test individual students, and to see that strategies in use match strategies stated in the proposal.

On an average of once a week, four bilingual program supervisors visit program teachers for whom they are responsible. During these visits, the supervisors check teachers' lesson plans, observe teachers in action, and discuss mutual concerns.

Reading specialists, mathematics specialists, and on-site coordinators visit bilingual teachers' classrooms as needed. In addition, the school district requires that each school principal observe all of his teachers in action at least twice a year.

Cost

Budgets and enrollments for the first five years of program operation were as follows:

Year	Enrollment	Title VII Budget	Total Budget (inc. other federal funds and local funds)	Title VII Per-Pupil Costs	Total Per-Pupil Costs
Year 1 1969-70	1,000	$197,927	$313,927	$197.93	$313.93
Year 2 1970-71	1,346	$438,092	$940,227	$325.48	$698.53
Year 3 1971-72	1,813	$518,413	$1,065,348	$285.94	$587.61
Year 4 1972-73	2,705	$527,013*	$1,269,276	$194.83	$469.23
Year 5 1973-74	2,830	$695,066*	$1,645,066	$245.61	$581.30

* Includes the Regional Curriculum Adaptation Center budget.

As can be seen in the five-year table, total monies allocated to the program have increased steadily. There are several reasons for the steady increase:

- The increased per-pupil cost in Year 2 reflects the expansion of the Arriba component into two more high schools and the implementation of Model B.
- Teacher salaries rose in Year 3 due to a retroactive pay increase totalling $40,000 for project teachers.
- Implementation and maintenance of a Regional Curriculum Adaptation Center increased project costs beginning in Year 4.
- A strike early in the program's fifth year of operation brought teachers another pay increase.
- As a result of the strike also, class size was limited to a maximum enrollment of 35 students per class; it therefore required more teachers to cover the additional classes which were created to maintain the maximum enrollment ceiling.

The program's reliance on federal funds other than Title VII increased from $33,000 in Year 1 to $245,000 in Year 5. Local support for the program increased from $83,000 in Year 1 to $705,000 in Year 5--nearly half of the total program cost. Beginning in Year 4, the local school district paid the salaries of all Model A teachers and the majority of Model B teachers.

EVIDENCE OF EFFECTIVENESS

Introduction. Evaluators in the Philadelphia Bilingual Program emphasize formative evaluation. The program is evolving and program changes have been and are being made on the basis of the continuous feedback provided by the evaluation. For example, because of the 1973-74 evaluation, increased emphasis is now being placed on English gains by Spanish-dominant students.

The program is worthy of consideration as an exemplary one for two major reasons. First, evaluation results have been favorable to the program when assessment has been based on comparisons between program and nonprogram students. Most program objectives have been met or surpassed.

Second, the program evaluation design is replicable and has utility for improving process evaluation in other bilingual education programs. It should also be noted that Philadelphia is one of the few bilingual education projects reviewed by AIR that has classified all tested students on the basis of their language dominance, and has reported results in first and second language performance. Such a basic concept in evaluating bilingual education projects has been frequently overlooked.

Evaluator. The evaluation was conducted by Dr. Robert Offenberg and his staff. Dr. Offenberg is a research associate with the Philadelphia School District and is assigned as a full-time evaluator for the bilingual project.

Evaluation year. 1973-74.

Aspects addressed. The evidence of effectiveness summary addresses the following aspects of the program evaluation:

- Language development--Spanish Mother Tongue
  English Mother Tongue
  English as a Second Language
  Spanish as a Second Language
- Self-concept
- Drop-out rate
- Principals' and parents' perceptions

Language Development

Spanish Mother Tongue

Sampling. Spanish-dominant students in Models A and B, and elementary and junior high school students of the Arriba component in grades 4 through 8 were tested on Spanish instruments. Scores for all students who were present at the time of the testing and who completed all subtests are reported. Some subtests were administered on different days, and therefore some students may have been present only during part of the testing. Because of the many project-administered measures as well as district testing, make-ups were not arranged for students absent during the testing period.

Comparison. Treatment students are compared to a pre-program baseline. This baseline was established in 1968 by testing all Spanish-speaking students in the Philadelphia schools who were seen as needing special attention due to limited English-speaking ability. This comparison is seen as reasonable because the same type of criteria has been used for determining participation in the program and in the developing curriculum. Grades 2 through 8 were administered the same Spanish tests used presently by the program. The baseline and treatment groups are compared to rural Puerto Rican norms, since the majority of students entering the program are reported as coming from rural Puerto Rican areas.

Measures and data collection. Students in the baseline group and treatment groups, tested in May, 1968 and May, 1974 respectively, have been tested in Spanish on the following measures:

Grades 1-3-----Test de Destrezas Basicas en Lectura

Grades 4-6-----Prueba de Lectura

Grades 7-8-----Inter-American Series, Level IV

Results. Table 1 shows the results on the Spanish language tests administered to Model A, Model B, and Arriba students, as well as baseline group results when these were available. The following statistical tests were made on the data summarized in Table 1.

Grades 2 and 3: On the Test de Destrezas Basicas en Lectura administered to pupils in Model A and Model B in Grades 2 and 3, an analysis of variance was calculated that indicated a significant grade effect (Grade 3 higher, $p = < .001$), a significant program effect (Model A higher, $p = < .001$) and no significant interaction. Two planned orthogonal comparisons were made: (1) pupils in Models A and B in Grade 2 showed a significantly higher ($p = < .001$) mean than did the baseline group; (2) pupils in Models A and B in Grade 3 showed a significantly higher ($p = < .001$) mean than did the baseline group.

Grades 4 and 5: Scores on the Prueba de Lectura for Model A pupils in Grades 4 and 5 were compared with scores for baseline groups in an analysis of variance. There were significant program effects (Models A and B pupil means higher than baseline, $p = < .001$), significant grade effects (Grade 5 higher, $p = < .002$), and no significant interaction.

Grades 7 and 8: Scores on the Spanish version of the Inter-American Reading Test for Arriba students in Grades 7 and 8 were compared with pre-program baseline pupils in an analysis of variance summary. Program effects were significant ($p =<.05$, one tailed), indicating Arriba pupils out-performed the pre-program baseline group. Grade effects were also significant ($p =<.01$) in favor of the Grade 8 students, and no significant interaction was revealed.

English Mother Tongue

Sampling. English-dominant students in Model A in Grades K through 5 who

TABLE 1

1973-74 Language Development: Spanish Mother Tongue

Grade	Test	Model A				Model B				Arriba				Baseline			
		N	M	SD	%	N	M	SD	%	N	M	SD	%	N	M	SD	%
Grade 1	Test de Destrezas Basicas	103	68	20	95	18	58	13	86	-	-	-	-	-	-	-	-
Grade 2	Test de Destrezas Basicas	78	89	16	75	39	81	22	62	-	-	-	-	266	58	21	35
Grade 3	Test de Destrezas Basicas	83	94	18	61	24	83	22	45	-	-	-	-	332	70	58[a]	27
Grade 4	Prueba de Lectura	46	26	16	25	9	22	10	13	15	20	12	10	180	18	9	07
Grade 5	Prueba de Lectura	22	28	17	18	-	-	-	-	15	16	13	01	192	21	10	07
Grade 6	Prueba de Lectura	-	-	-	-	-	-	-	-	16	45	13	36	-	-	-	-
Grade 7	Inter-American Reading	-	-	-	-	-	-	-	-	42	24	10	18	98	20	12	09
Grade 8	Inter-American Reading	-	-	-	-	-	-	-	-	36	27	13	14	82	24	15	08

[a] Program evaluator indicated some very high-scoring pupils increased the standard deviation for this group.

were present during the time of the testing and who completed all subtests were tested. Because of the many project-administered measures, as well as district testing, no make-ups were arranged for students absent during the testing period. Model A is the only one of the three models systematically serving English-dominant students.

Comparison. Treatment students are compared to a pre-program baseline. Since the measures used are the same as are used in the district testing, the baseline group is comprised of students at Potter-Thomas School who were at the given grade level the year before the program was cycled upward to that grade. Grade equivalents are used for comparison purposes.

Multiple analysis of variance was used to compare Model A pupils' scores on each test battery with pre-program baseline students. For each test battery, the vector of mean subtest scores for Model A pupils is compared with the vector of mean subtest scores for the pre-program baseline group. Also reported are univariate F-tests of group mean differences for each subtest in each battery.

Measures and data collection. Baseline groups and subsequent treatment groups have been tested in the Spring on the following measures:

Grade 2 : Stanford Achievement Test, Primary Battery II

Grade 3-5: Iowa Test of Basic Skills

Results. Table 2 presents means and standard deviations for Model A and baseline pupils in Grades 2, 3, 4, and 5 on the English language test administered.

TABLE 2

1973-74 Language Development: English Mother Tongue

			Stanford Achievement Test Primary II (Raw Scores)							
			Word Meaning		Paragraph Meaning		Spelling		Word Study Skills	
		N	M	SD	M	SD	M	SD	M	SD
Grade 2	Model A	96	11.6	5.8	15.7	8.9	7.0	6.1	25.5	9.4
	Baseline	42	6.1	3.8	8.6	6.3	2.1	2.1	16.5	6.3

			Iowa Test of Basic Skills (Grade Equivalents)					
			Vocabulary		Reading		Language	
		N	M	SD	M	SD	M	SD
Grade 3	Model A	88	2.6	0.7	2.6	0.9	2.8	0.7
	Baseline	89	2.3	0.6	2.3	0.6	2.3	0.5
Grade 4	Model A	56	3.5	1.0	3.4	1.0	3.5	1.0
	Baseline	74	3.4	1.4	3.2	1.1	3.4	1.0
Grade 5	Model A	33	5.3	1.3	4.9	1.3	4.1	1.3
	Baseline	56	3.3	1.1	3.4	1.0	3.4	1.0

Grade 2: The MANOVA overall test revealed significant mean subtest differences between Model A and baseline pupils on the Stanford Achievement Test Primary II (p = < .001). Also, univariate F-tests for each subtest showed mean differences favoring the Model A pupils for all subtests (p = < .001).

Grades 3, 4, and 5: The MANOVA summary indicated significant grade norm effects (p = < .001), significant mean vector differences favoring the Model A over the baseline groups (p = < .001), and also a significant interaction of Grade X Program (p =<.001). In every grade, Model A pupils showed higher grade equivalents than did baseline pupils on each subtest. Univariate F-tests also indicated significant mean differences favoring program pupils on each subtest. (It is noted that this test of program effectiveness is the comparison of means for all Model A pupils in Grades 3, 4, and 5 versus the means of baseline pupils in Grades 3, 4, and 5.)

English as the Second Language

Sampling. Spanish-dominant students in Models A and B and the Arriba component in Grades 3, 4, and 5 who were present on the day of the testing were included in the sample. In Model A, students in the regular English as a Second Language program were included. (Students new to the program for whom the emphasis was oral second language development were not tested.) In Model B and Arriba, all Spanish-dominant students who had "any experience" in reading English were included.

Comparison. Spanish-dominant treatment students are compared to the norms for English-speaking students. The expectancy was that Model A and B third grade students would perform two years below grade level--at a grade equivalent of 1.8--and that fourth and fifth grade Model A and B students would grow at a rate of one year grade equivalent for each grade level (2.8 and 3.8 respectively). No clear-cut expectancy was set for students in the Arriba component since they were recent in-migrants who had not had continuous education in Philadelphia schools.

Measures. Four subtests of the Stanford Achievement Test, Primary Battery II were administered: Word Meaning, Paragraph Meaning, Spelling and Word Study Skills.

Results. Table 3 presents the summaries by grade level on the English as a Second Language component. Results were mixed when compared with expected performance. In third grade, Model A achieved the criterion level and Model B did not. In fourth grade, Model B achieved the criterion level and Model A did not. In fifth grade, Model A did not reach the criterion level and Model B had not yet been cycled up to the fifth grade level. (It should be pointed out that the fifth grade group is the "pioneer group" in Model A, with 1973-74 being the first year that the project operated at that level.)

TABLE 3

1973-74 English Performance of Spanish-Dominant Pupils on Stanford Primary Battery II

(Grade Equivalents)

Group	Word Meaning			Paragraph Meaning			Spelling			Word Study Skills			Composite
	N	Mean GE	SD	N	Mean GE	SD	N	Mean GE	SD	N	Mean GE	SD	Mean GE
Model A													
3rd Grade	70	2.1	.51	71	1.9	.47	72	1.8	.64	70	2.0	.74	1.9
4th Grade	54	2.4	.83	54	2.1	.77	58	2.3	.11	57	2.4	1.17	2.3
5th Grade	29	2.4	.57	30	2.0	.41	28	2.2	.83	29	2.5	.84	2.3
Model B													
3rd Grade (Ludlow)	24	1.6	.34	21	1.7	.28	19	1.5	.31	15	1.1	.18	1.5
4th Grade (Ludlow)	10	3.0	.72	11	2.5	.70	11	3.2	.92	11	3.2	1.53	3.0
Arriba													
3rd Grade (Waring)	13	1.8	.18	14	1.8	.19	14	1.7	.59	15	1.6	.46	1.7
4th Grade (War&Lud)	14	2.3	.39	21	1.7	.20	18	1.8	.52	18	1.8	.42	1.9
5th Grade (Lud&Swrk)	21	2.0	.59	24	2.0	.62	26	1.8	.87	26	1.9	.62	1.9
All Components													
3rd Grade	107	1.9		106	1.9		105	1.7		100	1.8		1.8
4th Grade	85	2.4		86	2.1		87	2.3		86	2.4		2.3
5th Grade	50	2.2		54	2.0		54	2.0		55	2.2		2.1

Spanish as a Second Language (SSL)

Sampling. All English-dominant students in Model A participating in regular SSL instruction in fourth and fifth grades were tested. New students to the program for whom the emphasis was oral second language development were excluded. All students present on the days of the test administration were examined.

Measures. Test de Destrezas Basicas en Lectura, designed for first through third grade native Spanish-speaking students.

Comparison. SSL students are compared to rural Puerto Rican, third grade, end-of-the-year norms.

Results. Compared to rural Puerto Rican, end-of-year, third-grade norms, English-dominant students at both fourth and fifth grade levels were performing near or above the norms. The Letter and Word Recognition subtest requires more knowledge of phonics than of Spanish language, while the Word Meaning and Paragraph Meaning subtests do require knowledge of the Spanish language.

Table 4 presents the means, standard deviations, and percentile ranks on the Test de Destrezas Basicas en Lectura for pupils in grades 4 and 5.

TABLE 4

Performance of English-Dominant Fourth- and Fifth- Grade Pupils On Spanish Reading Test, Test De Destrezas Basicas En Lectura

Item	Letter and Word Recognition				Word Meaning				Paragraph Meaning				Composite			
	N	$\bar{X}$	SD	PR*	N	$\bar{X}$	SD	PR	N	$\bar{X}$	SD	PR	N	$\bar{X}$	SD	PR
Grade 4	78	60	5	80	76	14	6	40	83	13	5	50	74	87	13	60
Grade 5	32	62	5	90	32	16	7	50	31	16	5	66	30	95	15	70

* Percentile rank is based on rural norms of Puerto Rico, third grade end-of-the-year.

Self-Concept

Sampling. 166 Spanish-dominant students were randomly selected from the Model A and Arriba components in fourth through twelfth grades.

Comparison. Bilingual treatment students are compared to 100 randomly selected students participating in the district English as a Second Language program in schools where there is no regular bilingual program. For analysis purposes, the students were divided into three groups: elementary (grades 4-6), junior high school (grades 7-9), and high school (grades 10-12).

Measure and data collection. The instrument used was a translation of the Coopersmith Self-Esteem Inventory. The instrument was translated by a project staff member and field tested and revised the summer prior to its use. It was orally administered to small groups by a bilingual member of the project evaluation staff.

Results. The analysis shows that there were statistically significant differences between the programs and between the grade levels and a significant interaction between these two variables. At the elementary and junior high school levels, pupil self-esteem is similar. However, there is a strong difference between the two groups at the senior high school level, with the bilingual project participants clearly demonstrating a higher level of self-esteem. Results are graphically illustrated in Figure 1.

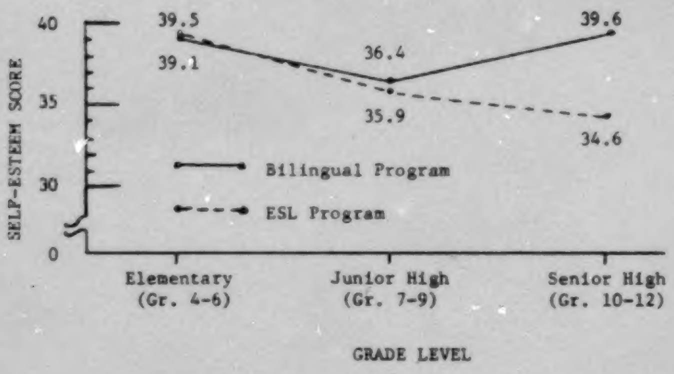

Figure 1. Comparison of self-esteem scores of bilingual program and ESL-only groups.

Drop-Out Rate

Sampling. All bilingual program students in the Arriba component who were in the tenth grade in December, 1971 and who were classified as Spanish speaking in the city-wide pupil directory were included in the sample (N-104).

Comparison. The bilingual treatment sample was compared to all other students in the treatment schools, and city-wide, who were listed in the city-wide pupil directory as Spanish speaking but who were not participating in the bilingual project. The comparison group consisted of 264 students not in Arriba, but in the same schools, and 451 students city-wide. The comparison consisted of the percentage of students enrolled in December, 1971, who were graduated in June, 1974. This longitudinal approach was taken to control for summer drop-outs between grade levels. (Within-scholastic year analyses done in previous years have also shown consistently favorable results.) Results are shown in Table 5.

TABLE 5

Graduation Rate of Spanish-Dominant Students Who Were in Grade 10 in December 1971

Group	Graduated		Not Graduated		Total	
	N	%	N	%	N	%
In Arriba	30	29%	74	71%	104	100%
Not in Arriba (same schools)	52	20%	212	80%	264	100%
Not in Arriba (city-wide)	115	26%	336	75%	451	100%

Results. Bilingual project students were more likely to graduate than other Spanish-dominant students in the same schools. The difference was statistically significant at the .05 level of probability. When compared to city-wide Spanish-dominant students, the difference was not statistically significant, but favored the bilingual project. (See Table 5.)

Opinion Surveys

Summary. Opinion surveys of principals of project schools and parents of participating students were conducted in May, 1974. Principals in the eleven participating schools were asked for their general, overall impression of the project and for specific information about pupil and parent reactions to it. Parents were asked for their perception of the program and about their participation in and contact with the project.

Sample. Principals: All eleven principals of participating schools responded.

Parents: A total of 677 parents responded, or 34% of the parents of participating students.

Measures. Principals were mailed questionnaires in May, 1974. The parents' questionnaire was in English and Spanish. It was sent to the parents with all the students who were present on the day the questionnaire was distributed. Since the questionnaire was anonymous, there was no way of following up on those which were not returned.

Results for the Principal Questionnaire. The questionnaire consisted of seven questions. The first two questions were general, dealing with overall satisfaction with the project and opinions about expansion. Questions 3 through 6 dealt with teacher supervision, the summer institute, and increased understanding of the project by parents, students, and staff. Question 7 asked for additional areas of concern that principals wanted to bring to the attention of project management.

Dealing with the two general questions about the project, the responses were as follows:

a. Overall level of satisfaction with the bilingual project: 5 were satisfied, 6 were somewhat satisfied; none were somewhat dissatisfied or very dissatisfied.

b. Opinions about expanding the program within their schools: 4 expressed desire that the program be expanded to reach more students; 7 felt that the program should remain the same size (5 responded, "already served children who need it," 3 responded, "limitations in terms of space and teacher vacancies to allow employment of bilingual teachers"; no one felt the program should be reduced or eliminated.

Results for the Parent Questionnaire. Although caution must be exercised in viewing these data because of self-selection of the respondents (34% response rate), the results indicate a high level of support for the program. The questionnaire is divided into three sections: (a) home language use, (b) perceptions of school program, and (c) parent participation.

On three questions dealing with perception of the school program, the parents responded in the following manner:

a. 93% of the parents felt that their children liked learning Spanish and English in school.

b. 94% of the parents liked their children to be learning Spanish and English in school.

c. 93% of the parents wanted their children to study two languages the following year.

PROGRAM AREA:

Bilingual Education (Spanish/English)

PROJECT TITLE:

Alice Independent School District Bilingual Education Program

LOCATION:

Alice, Texas

SOURCE AND LEVEL OF FUNDING:

The following figures are for the fourth year of operation.

Title VII Funds	$100,625
Other Federal Funds	75,820
State and Local Funds	201,167
Total	$377,612

PROGRAM START DATE:

1970

BRIEF DESCRIPTION OF PROJECT:

Goals and objectives. It is the goal of the project to teach children with limited English-speaking ability to speak English while also developing proficiency in Spanish. It is felt that developing both languages will facilitate achievement in school. It is also the intent to develop and maintain the children's pride in themselves and both cultures and to involve parents from non-English-speaking environments in school-related activities.

Context. Alice is a small, semi-industrial town of about 20,000 people. It is 40 miles west of Corpus Christi and 120 miles north of the U.S./Mexican border town of Hidalgo, Texas. Much of the employment is in services related to oil mining. The population of Alice is very stable, with the exception of migrant workers, many of whom return to Alice each year after their field work is finished. There is some mobility within the town, but attrition from the program has not been a problem.

The total enrollment in the Alice Independent School District is 6,268. About 69% are Mexican-American, and 31% are Anglo-American. It is estimated that 66% of the total enrollment has limited facility in the English language. The proportion of Mexican-Americans in the four elementary schools offering bilingual education in the 1973-74 year was about 96%.

Major features of the program are the following:

- the use of performance objectives for all subjects taught bilingually,

- teacher participation in determining certain elements of the curricula,
- comprehensive record-keeping,
- flexibility of instructional approach,
- intensive preservice and inservice training, and
- active parent and community involvement.

These features will be discussed in appropriate sections below.

Program description.

Years of operation, size, grade levels--In 1973-74, there were 528 children in grades K-4 in four schools. This year the program has expanded horizontally into all nine elementary schools, two of which are parochial. It has expanded vertically into the 5th grade.

Staff--Staff members who were full time on the bilingual project in 1973-74 were the director, an evaluator/materials specialist, 19 teachers, 6 teacher aides, and a secretary. All central staff are bilingual. There were also 7 mother aides who had children in the program and who worked part time. Staff from other programs administered by the district (i.e., Title I, Migrant Education, Career Education, Community Involvement) freely share their resources and experience with the Title VII staff.

Curricula and time involved--Children generally enter the program in kindergarten. They are preregistered by their parents in the spring before they enter kindergarten and are accepted on a "first-come, first-serve" basis. A language proficiency test, provided by the Texas Education Agency, is administered in English first and then in Spanish to determine the child's proficiency in each language.

The instructional component for all grade levels centers around performance objectives in the following content areas: language arts (including listening, speaking, reading readiness, and reading skills), mathematics, culture and heritage, and science. The objectives, which were developed by program staff, are written in behavioral terms. Mastery tests for each objective have also been developed. For each objective at every grade level, there is a designation as to which language(s) the objective should be mastered in and what percentage of students are expected to master it. The objectives are revised every summer by teachers. During the year, if teachers discover that an objective is too easy, too difficult, or somehow inappropriate, they consult with the director and evaluator and either drop the objective or rewrite it.

Teachers submit lesson plans to the program director every two weeks that indicate which objectives are being taught, in which language they are being taught, and types of activities planned. The director reviews these lesson plans and returns them to the teachers.

Detailed records are maintained on each child in the program indicating how many objectives he has mastered and his score on each mastery test. In addition to keeping a class roster, teachers maintain a separate record for each child, which becomes part of his permanent file. This system enables teachers to know at the beginning of each year exactly at what level each child is performing. Records are spot-checked throughout the year by the director.

Instruction begins in kindergarten in Spanish in reading readiness, Spanish language arts, and mathematics. Other activities such as group time in the morning, story time, music, art, and afternoon review are also conducted in Spanish. Language arts in English concentrates on developing oral language skills and in highlighting the similarities and differences in the sounds and alphabets of both languages. The concepts taught in Spanish are reinforced in English. As kindergartners become proficient in Spanish and begin to develop basic skills in English, more time is devoted to instruction in English. Kindergarten children also view an early intervention program for preschool children carried on local cable television. The program consists of a 30-minute color broadcast each day produced at the Alice High School. It is bilingual and is designed to upgrade the cognitive and affective skills of preschool children in Alice. Concepts presented during the first 15-minute segment are reinforced with a story during the second half of the program.

Reading instruction begins in grade 1 in Spanish by a "building block" approach. Children learn the sound of each letter in the Spanish alphabet, its name, and how it looks. They then put sounds together to make syllables, and put syllables together to make words. They also begin manuscript writing in grade 1. English vocabulary and language skills continue to be reinforced. In late fall or early winter, the reading readiness test that accompanies the district reading series is administered to small groups of first graders. Based on results of this test and teacher observation, reading instruction begins in English. For all children this is by January of grade 1. As the child's performance in Spanish improves and his proficiency in English develops, equal time is spent instructing in both languages in language arts, math, and culture and heritage. For most children, this is during grade 1.

In grades 2 through 5, instruction continues to be given in English and in Spanish for equal amounts of time. Some teachers teach one week in Spanish and the following in English; others alternate every other day. Reading in Spanish and in English is taught daily. The program provides teachers a great deal of flexibility, and they are free to use their own judgments about when is the best time to move into completely bilingual instruction.

At all grade levels, program classrooms are heterogenous and self-contained, and teaching methods vary. Grouping is generally informal and is determined by language proficiency and/or achievement level. There is a shifting emphasis to greater individualizing of instruction, although this approach is not imposed as protocol on teachers. Some teachers use interest centers as a means of providing a variety of learning experiences; others employ more traditional techniques.

Materials--The following is a summary of key materials used at all grade levels.

> Reading readiness skills are taught in Spanish and English using Peabody Language Development Kits, ROCK materials developed by the Region One Education Service Center in Edinburg, Texas, The Alphabet Book and Really Reading Series published by Harper and Row, and Alpha Time with the Huggables by New Dimensions in Education. Teachers also have access to teaching machines, language kits, and other audiovisual equipment.

Reading instruction in Spanish is taught using Metodo Onomatopeyico, Mi Libro Magico from the Continental Series and Preparandose Para Leer by Houghton-Mifflin. The Laidlaw Series, Por El Mundo Del Cuento Y La Aventura, is the core of the Spanish reading program in grades 2-5.

Reading in English is taught with the district-adopted Harper Row Reading Series.

Language arts are taught in Spanish with Ejercicios de Lenguaje, a grammar book from Mexico, and Our Language Today from the American Book Company

Supplementary materials for each reading series (workbooks, flashcards, charts, etc.), staff-developed readers, and many other library books in Spanish and English enrich the reading and language arts programs.

For mathematics instruction, The Elementary School Mathematics Series by Addison-Wesley is used throughout the district.

Science is taught bilingually in grades 4 and 5 using the McMillan text, Science for Tomorrow's World. The staff has written units in Spanish to supplement the text.

The staff has also developed a program to teach culture and heritage called Mi Herencia Cultural, which is adaptable to any grade level. Some of the topics included are the immigration of Mexicans to America, the migration of Mexican-American farm workers, and recognition of prominent Mexican-Americans in Alice.

Facilities--The program operates in regular classrooms. No modifications of facilities are required.

Preservice/inservice training--In August, all Title VII staff members attend a comprehensive preservice training session. During the summer 1973 session, which involved 80 hours' training, consultants in linguistics and South Texas Mexican literature spoke to the staff. Representatives from publishers also demonstrated new materials. Teachers revised performance objectives and wrote mastery tests for all grade levels. They also developed instructional materials, which were evaluated by two consultants from a nearby university. Staff also discussed problems encountered the previous year and sought solutions to them. Teachers and aides receive stipends for attending the workshops.

During the school year, 10 days are set aside for inservice training. For teachers, programs center around educational television, teacher attitudes, teacher assessment, and a review of performance objectives. Topics for teacher and mother aides reflect the project's commitment to involve assistants totally in the instructional process. Some of their training includes learning how to listen to recitations, work with small groups, assist with drill exercises, and work on special cultural projects with students.

Teacher performance is videotaped twice a year using the Flanders Interaction Analysis, a system designed to measure how much emphasis the teacher places on the content of what is being presented, on students' ideas, and on the use of authority. The director reviews each teacher's tape with the teacher.

Other staff development activities are visits to other bilingual education programs throughout the state and attendance at conferences in other parts of the country. Teachers are also working toward receiving certification in bilingual education, which is soon to be mandatory in Texas. Certification involves taking a 30-hour class given by one of the regional offices, at least one year teaching experience, and demonstrated proficiency in Spanish.

Parent/community involvement--To encourage parents to become active participants in their children's education, priority has been given to parent involvement in the project. The Title VII Advisory Committee meets every six weeks on different school campuses. Membership is open to anyone interested in bilingual education. A constitution was written and approved during 1973-74 and has contributed to improving the Committee's status as an organization. There are five officers. Throughout the year the Committee has staged children's programs at their meetings, has met with members of the Title I advisory group, and has helped teachers present a Mother's Day program. The Committee's approval of new proposals is also solicited.

Parents volunteer to work in the school libraries and have prepared many of the consumable materials used in classrooms. Workshops are held where they learn how to care for and repair library books and how to perform certain jobs in the library. Encouraging parents to work in the library is seen as one way to stress the importance of reading--an emphasis that hopefully they will convey to the children.

Workshops, generally scheduled by the Home Coordinator, are also held to teach parents how to be effective learning aids in school and at home. Parents are welcome in the classrooms at any time and are willing volunteers when extra help is needed, such as on field trips or preparing for a special program. Two schools have Parent Involvement Centers, equipped with a television, sofas, tables, chairs, shelves, and books. Parent-teacher conferences are held twice a year to discuss how the child is proceeding through the program.

Dissemination of information about the project to the public is another essential element of this component. Some of the resources utilized to keep the community informed of the bilingual events are newspaper articles, newsletters in Spanish and English, Christmas cards, and announcements in the paper about every event in the project of public concern.

Costs--Total budgets and enrollments for the first four years of program operation were as follows:

	Enrollment	Total Budget	Per Pupil Cost
Year 1	150	$101,150	$674
Year 2	270	$285,150	$1,056
Year 3	3[illegible]	$366,730	$1,095
Year 4	528	$377,612	$715

The per pupil cost is computed on the total cost of the program, including teacher salaries, and is the amount expended for the total educational program for each child in the bilingual education program. The per pupil cost for the regular district program in Alice was $764 in 1972-73 (Year 3 of the bilingual program) and $785 in 1973-74 (Year 4) and includes all expenditures except bonded indebtedness and construction.

Major start-up costs in Year 1 were for instructional materials and consumable supplies ($12,646), capital outlay for instructional equipment, furniture, and non-consumable instructional materials ($9,634), contracted services ($2,000), and travel for administrative and teaching staff ($4,780). In general, the amounts spent for major start-up expense categories have dropped, except, of course, for staff salaries and instructional materials as the program expands to include more classes and additional grade levels each year. A significant rise in per-pupil costs occurred in Year 2 and Year 3, when this figure was about $1,000 per child. This increase was incurred when teachers and paraprofessionals began to be paid for participating in the intensive preservice training regimen. These stipends are paid to all participating teachers and paraprofessionals and amounted to $3,690 in 1971, $5,700 in 1972 (Year 3), and $6,880 in 1973 (Year 4).

Also contributing to the rise in per-pupil cost in Year 2 was the addition of paid mother aides. In year 4, the per-pupil cost dropped back closer to the Year 1 figure, about $715 per child. The reason for the drop, according to program staff, was that the amount spent for budget items from previous years did not increase proportionately with the expanding program enrollment. For instance, the amount spent on materials, equipment, and supplies was sharply reduced so that the program could continue with local funds should Title VII monies cease to be available. Program staff found that materials could be catalogued so that teachers could share what had already been purchased without detriment to the instructional program.

EVIDENCE OF EFFECTIVENESS:

Overview. During the present 1974-75 academic year, the bilingual education program in the Alice Independent School District (AISD) is in operation in 28 classes in nine schools. These are Saenz, Nayer, Salazar, Garcia, Noonan, Hillcrest and Schallert in the public school system, and St. Joseph's and St. Elizabeth's in the parochial system. The first four schools were those that had been continuously involved in the bilingual program at AISD for five years since its inception in 1970. Since its start, the program has grown both horizontally (more classes at a given grade level) and vertically (expansion through higher grades). Table 1 shows the extent of the program's growth at AISD both in terms of total number of classes and in the number of participants each year.

TABLE 1

Classes Within Grade Levels With Schools at
Alice Independent School District

	Bilingual Classes for Academic Year				
Schools	1970-71	1971-72	1972-73	1973-74	1974-75
*Saenz	1	1,2	K,1,2,3	K-4	K-5
*Nayer	1	1,2	K,1,2,3	K-4	K-5
*Salazar	K	K,1	K,1,2	K-3	K-4
*Garcia	1,1	K,1,2	K,1,2,3	K-4	K-5
Noonan					1
Hillcrest					1
Schallert					1
St. Joseph's					1
St. Elizabeth's					1
Total Classes	5	9	15	19	28
Total Students	150	270	417	528	766

*Title I schools

The evidence of effectiveness of the AISD bilingual program is reported for the third and fourth year of the program's operation; pre- and posttest scores and gain scores on the Inter-American Series Reading (Lectura is the Spanish version title) and General Ability (Habilidad General is the Spanish version title) tests are compared for program and control pupils enrolled in kindergarten through the third grade in 1972-73 and in kindergarten through the fourth grade in 1973-74.

Evaluation. The 1972-73 year's evaluation of the AISD bilingual program was made by V. J. Kennedy, Educational Program Auditor, University of Houston; the 1973-74 year's evaluation was made by Arnoldo Salinas, Project Director, and Claudina Hernrndez, Project Evaluator.

Year and grade samples. Table 2 shows the number of program and control group participants with complete pre- and posttest data on the Inter-American Tests.

TABLE 2

Bilingual Program and Control Pupils
With Complete Pre- and Posttest Data

Grade	1972-1973		1973-1974	
	Bilingual	Control	Bilingual	Control
K	124	32	127	33
1	104	27	106	19
2	94	21	95	35
3	75	22	101	29
4	___	___	75	20
TOTALS	397	102	504	136

Bilingual group--Children who enter kindergarten in the Alice Independent School District are tested with the Oral Language Eligibility Test for Kindergarten Pupils. This is an individually administered test to determine oral English ability based upon the quality of verbal responses to 20 questions. Children performing poorly on this test and who are registered for the bilingual kindergarten are placed in the bilingual class if space is available. Bilingual classes never exceed 35 pupils so that participation in kindergarten is a function of test score, parental approval and space. New children entering elementary grades who experience English language problems and whose parents desire their enrollment in a bilingual program are placed in a bilingual class. (In 1972-73, one third-grade bilingual class was reduced in size because of a residential move by several parents to a newly constructed housing development. This class was increased by replacements whose parents agreed to their participation when requested. These replacement children were average or above in ability since they would need to acquire Spanish language skills quickly; the group of children who left the program at this time and their replacements were both Spanish dominant.)

Control group--There is one control class at each grade level. The project director attempts to pick a regular class whose pupils' Oral Language Eligibility Test scores match the bilingual kindergarten class's most closely; this is then the control class for that year's entering kindergarten group. All control classes are located in Title I schools. This control class is essentially a control cohort as it progresses through grades K, 1, 2, and so on. New children entering elementary grades placed with the control cohort join it if both pre- and posttest data are available. So far, no control student has later become a bilingual program student, and no bilingual program student a control student later.

Measures.

Comparison--Pupils in the bilingual and control classes are administered both the English and Spanish versions of the Reading and General Ability Tests of the Inter-American Series twice a year. Pretesting is during the first two weeks in September; posttesting in mid-to-late April. This retest interval is approximately seven months.

For each language version, different forms of the tests are administered at pre- and posttesting. The control groups at each testing are first administered the Spanish version, followed the next day by the English version. Bilingual pupils first take the English version, then the Spanish. All testing takes place in the morning. Table 3 summarizes the general procedures followed by AISD for administering the Inter-American Series at each grade level.

TABLE 3

Order of Test Administration

Group	Pretest 1	Pretest 2	Posttest[a] 1	Posttest[a] 2
Bilingual	English	Spanish	English	Spanish
Control	Spanish	English	Spanish	English

[a] In every case a different form of the pretest is administered as a posttest.

Other tests--Beginning with the 1974-75 academic year, plans were implemented to also administer the Reading and Mathematics SRA Achievement Series to all pupils in the elementary grades from grade 2 upward in AISD. Therefore the bilingual and control groups in grades 2-5 in 1974-75 have been pretested with the SRA Achievement Series and are currently being posttested with different forms; comparative gain data, as well as comparative data for the bilingual and the entire mainstream group will be available in the summer of 1975.

Mastery tests associated with the Harper and Row Reading Series are administered, but no comparative basal reading data summaries for the bilingual and regular students are prepared.

Comparison procedures. For both the 1972-73 and 1973-74 school years, mean pretest, mean posttest, and mean gain (posttest minus pretest) on the Inter-American Series Reading and General Ability Tests are reported for the control group and for the combined bilingual classes at each grade level.

Mean pretest, posttest and gain scores for the control and the bilingual program groups enrolled in K through 3 in 1972-73 and K through 4 in 1973-74 are shown in Table 4.

TABLE 4

Pretest, Posttest and Gain Means for Bilingual Program and Control Pupils Classified by Grade Level [1]

Inter-American Series Test

				English						Spanish					
				Reading			General Ability			Lectura			Habilidad General		
Year	Class	Group	N	Pre	Post	Gain	Pre	Post	Gain	Pre	Post	Gain	Pre	Post	Gain
1972-73	K	Bilingual	124	--	--	--	48.05	67.72	19.67	--	--	--	57.30	71.06	13.76*
		Control	32	--	--	--	49.06	70.74	21.68	--	--	--	42.41	65.56	23.15
	1	Bilingual	104	23.27	43.76	20.49*	48.26	69.03	20.77	23.44	57.07	33.63*	52.57	74.01	21.44*
		Control	27	17.55	30.89	13.34	40.10	62.11	22.01	13.71	24.93	11.22	47.44	63.69	16.52
	2	Bilingual	94	38.31	62.33	24.02	42.47	61.35	13.88	45.20	64.68	17.98*	50.90	68.82	17.92*
		Control	21	30.33	51.24	20.92	46.39	55.86	9.17	22.89	27.67	4.78	44.56	48.14	3.58
	3	Bilingual	75	41.50	64.44	22.94	56.44	67.53	11.09	43.43	60.30	16.87*	59.68	69.01	9.33
		Control	22	48.60	68.05	18.45	61.64	65.55	3.91	32.40	37.91	5.51	59.28	58.59	0.69
1973-74	K	Bilingual	127	--	--	--	48.62	65.44	16.82	--	--	--	54.87	69.95	15.08*
		Control	33	--	--	--	52.59	64.64	12.05	--	--	--	47.94	57.30	9.36
	1	Bilingual	106	23.06	45.86	22.80*	52.86	71.46	18.60*	25.14	57.50	32.36*	55.95	72.28	16.33*
		Control	19	20.67	36.05	15.38	47.83	58.26	10.43	20.22	20.05	-0.17	42.33	48.63	6.30
	2	Bilingual	95	38.23	61.01	22.78	49.87	70.93	21.06	45.89	63.54	17.56	53.71	75.01	21.30
		Control	35	25.81	43.80	17.99	44.72	59.31	14.59	20.34	32.74	12.40	40.69	59.74	19.05
	3	Bilingual	101	51.95	67.08	22.13*	58.13	74.48	16.35	47.62	64.99	17.37*	61.08	74.71	13.63
		Control	29	44.49	55.66	11.17	58.23	67.00	8.77	27.94	30.86	2.92	48.89	61.03	12.14
	4	Bilingual	75	54.95	77.87	22.90*	56.18	75.01	18.83	58.25	78.95	20.70*	60.58	81.09	20.51*
		Control	20	71.10	72.45	1.35	61.65	80.18	18.53	40.60	37.45	3.15	56.45	62.86	6.41

*Significant gains (see page 11).

[1] In general, bilingual project of control pupils in a particular grade level were also bilingual project or control students, respectively,the year before. Thus while there is no difference in the English General Ability pretest means for the two kindergarten groups (48.05, 49.06) in 1972-73, there is a pretest means difference favoring the 1973-74 first grade bilingual group (52.86, 47.83).

No significance tests of mean gain difference have been reported although the project evaluation reports refer to statistically significant differences in gains between project and control students. The raw data summaries are unavailable. However, it is possible to conservatively estimate the significance of the mean gain differences between the two groups. The following bases and assumptions are made:

- Posttest standard deviations are reported for classroom groups at each grade level. Pooled estimates for the bilingual grade groups are K = 7.0, 1 = 7.0, 2 = 12.0, 3 = 15.0, 4 = 20.0. Pretest standard deviations will be assumed to be the same; this is a conservative assumption since posttests nearly always show greater variability than do pretests.

- There is no estimate for the correlation between pre- and posttests. It is assumed to be 0.00; this is a conservative assumption since pre- and posttests in heterogenous groups are usually positively correlated.

- The standard error of mean gain differences is equal to

$$\sqrt{\frac{1}{N(\text{control})} + \frac{1}{N(\text{project})}} \times \text{SD gain (pooled)}$$

$$\doteq .25 \sqrt{SD^2\text{pre} + SD^2\text{post} - 2r \text{ pre-post SD pre SD post}}$$

$$\doteq .25 \sqrt{2} \text{ SD post (pooled)}$$

$$\doteq .35 \text{ SD post}$$

- A significant gain difference would be about twice its estimated standard error or .70 SD post. Based on the above estimates, significant differences are K = 5 points, 1 = 5 points, 2 = 9 points, 3 = 11 points, and 4 = 14 points.

Significant results.

1973-74 groups--Every mean difference favors the bilingual program group in the 1973-74 comparisons (see Table 4). Conservative estimates of the significance of the mean gain difference between the project and control groups show "significant" differences favoring the project groups in grades 1, 3 and 4 in English Reading, in grade 1 in English General Ability, in grades K, 1, 3 and 4 in Spanish Reading and in grades 1 and 4 in Spanish General Ability.

In terms of educational significance, mean gain differences are approximately equal to one standard deviation (posttest) in grades K and 1 on all four tests. In grade 2 the mean differences were at least one-half of a standard deviation unit (posttest) except for Spanish General Ability. In grade 3, mean differences were at least three-quarters of a standard deviation unit except for Spanish General Ability. In grade 4, the mean differences were at least one standard deviation unit except for English General Ability.

1972-73 groups--All but three mean gain differences favor the bilingual program group. The exceptions are the two versions of the General Ability Tests in K, and the English General Ability Test in 1. "Significant" mean gain differences favoring the program groups were found in English Reading in grade 1, Spanish Reading in grades 1, 2, and 3, and in Spanish General Ability in grades 1 and 2.

As for the educational significance of the comparisons, the mean gain differences were approximately one standard deviation (posttest) or higher on all tests in grade 1 except for English General Ability. In grade 2 mean gain differences were approximately one-half of a standard deviation unit for the English tests and at least one standard deviation unit for the Spanish tests. In grade 3, mean gain differences were at least one-half a standard deviation unit except for English Reading.

Summary. Although exact statistical tests of significance are not available showing mean gain comparisions for program and control students, the mean differences with few exceptions favor the bilingual pupils. The size of the mean differences are of obvious educational significance; that is, there is considerable non-overlap in the gain distribution for the two groups.

The procedure for selecting controls at the kindergarten level is probably acceptable, given that no real matching seems possible. Controls should be better in oral English at the entering kindergarten level. Thereafter, program effect can be revealed in two ways. First, through increased gains during each school year for the program participants over the controls, and second, through superior posttest performance for the program pupils at each grade level. The data bear out both of these types of program effect indicators. Test administration order, with bilingual pupils administered the English version of the Inter-American Tests before the Spanish version, and with control pupils administered the Spanish version before the English version, would tend if anything to introduce practice effects into the Inter-American English version scores for the control students.

With respect to achievement gain comparisons, interpretation is clearer if the groups to be compared are equivalent at pretest. This is not possible in this situation. The controls are "most equivalent" (to the extent that is possible) at the kindergarten level. After that, mean pre- and posttest, as well as gain comparisons, overwhelmingly favor the program pupils as shown in Table 5.

TABLE 5

Mean Gain Comparisons Favoring the Bilingual Program Pupils

	1972-73			1973-74			
Grade	Pre	Post	Gain	Pre	Post	Gain	Overall
K	1 of 2	1 of 2	0 of 2	1 of 2	2 of 2	2 of 2	7 of 12
1	4 of 4	4 of 4	3 of 4	4 of 4	4 of 4	4 of 4	23 of 24
2	3 of 4	4 of 4	4 of 4	4 of 4	4 of 4	4 of 4	23 of 24
3	2 of 4	3 of 4	4 of 4	3 of 4	4 of 4	4 of 4	20 of 24
4	--	--	--	2 of 4	3 of 4	4 of 4	9 of 12
Totals	10 of 14	12 of 14	11 of 14	14 of 18	17 of 18	18 of 18	82 of 96

PROGRAM AREA:

Bilingual Education (Spanish/English)

PROJECT TITLE:

Bilingual Education Program (Aprendemos En Dos Idiomas)

LOCATION:

Corpus Christi, Texas

SOURCES AND LEVEL OF FUNDING:

The following amounts were allocated for the fourth year of operation:

Title VII	$ 91,000*
Title I	34,600
Local	17,202
Minimum Foundation Program	180,000**
Total Year 4 Allocation	$322,802

*Actual expenditures from the Title VII allocation were $82,202.
**This sum includes state and local monies for teacher salaries.

PROGRAM START DATE:

1970

BRIEF DESCRIPTION OF PROJECT:

Goals and objectives. The Corpus Christi Independent School District has aimed at providing an effective bilingual-bicultural program of instruction for children of Mexican-American descent. The District has chosen four ways of reaching that goal: (1) the use of a set of instructional strategies and routines selected specifically for the program; (2) the review, selection, adaptation, and/or development of appropriate materials; (3) the implementation of a staff development program tailored to program strategies and teacher needs, and (4) the direct, extensive involvement of parents through voluntary participation in a variety of program activities that depend on their help.

Context. Corpus Christi is a city of over 200,000 people on the coast of Texas. It is supported by agriculture, petrochemical and metal industries and manufacturing plants, but is also the mercantile center for much of the southern part of Texas. The Port of Corpus Christi, one of the nation's top 10 harbors, is just next to the downtown business district. The population is mobile, particularly the portion of Mexican-American descent.

In 1974, the school enrollment in the Corpus Christi Independent School District was about 44,000, of whom 54% were Spanish surnamed. In the project, however, 92% of the children have Spanish surnames. According to parents of participants, about 74% of the children are native Spanish speakers. A teacher survey in the three project schools revealed that about half of the children relied most on their native Spanish language for communication; the teacher opinions were verified by administering the Oral English Proficiency Test (Elizabeth Ott) to a 10% random sample of project population.

Among families whose children are in the project, 81% receive incomes within the poverty level. The average family has 4.5 children. The absentee rate for the district is high, particularly in kindergarten classes.

Major features of the program. Key program strategies which were evident from on-site observation and program documentation were as follows:

- Use of bilingual teachers exclusively at Kindergarten and Grade 1; where classes must be assigned to monolingual teachers at Grades 2 through 4, pairing them with bilingual teachers for cooperative teaching
- Providing teachers with weekly visits and personal help from bilingual Instructional Consultants who have classroom experience in bilingual education
- Daily, cooperative planning among teachers and aides by grade level
- Careful planning, coordination, and monitoring of the program by administrative staff
- Introduction of English reading readiness and aural-oral skills as soon as a child demonstrates minimal aural comprehension of English
- Emphasis on English reading and communication skills for the duration of the program
- Use of bilingual teacher aides for instructional support and reinforcement
- Use of teacher-evaluated and teacher-developed curriculum materials tailored to participants' needs
- Providing intensive preservice training for all project staff, reinforced by monthly inservice days during the school year
- Promoting parent commitment to the program through their direct involvement in many classroom-related activities

Each of these major program features will be explained more fully in the sections that follow.

Grade levels, years of operation, size. During its fourth year of operation, 1973-74, the program served 519 children in grades K through 3.

Staffing. The program staff required in 1973-74 for this number of participants was as follows: one full-time Project Administrator, one full-time Instructional Consultant/Parent-Community Involvement Coordinator, one full-time Evaluator/Staff Development Coordinator, 22 full-time teachers, and 15 full-time teacher aides. (Eight of the aides served the eight kindergarten teachers fulltime; the other seven aides divided their time among the 14 teachers in grades 1 through 3 classes.)

All staff were bilingual. In 1974-75, with the expansion of the program through Grade 4, three monolingual teachers were added to the staff to maintain an ethnic balance among instructors that matched the ethnic balance of the project schools. A second full-time Instructional Consultant was also hired. The Program Evaluator is now part-time in order to free salary funds for the second Instructional Consultant.

Staff roles in program management -- To help teachers implement the program properly, an Instructional Consultant visits every project classroom at the three participating schools at least once every week and gives instructional support in a variety of ways, e.g., procurement of instructional materials, assistance in working with individuals or groups, demonstration of teaching techniques with individuals or groups, advice on individualizing instruction. The Instructional Consultants have several years of experience teaching bilingually and feel this has helped them to be effective and to establish rapport with the children and their teachers, teacher aides, and parents.

The Program Administrator also visits classes routinely and informally, and drops in occasionally to join the teachers' planning periods. His special concern is maintaining good rapport between the teachers and the principals and between principals and program administrative staff.

Teachers at each grade level meet four days a week during the children's physical education period to coordinate lesson plans. Where monolingual and bilingual teachers trade groups for English and Spanish reading, language arts, and math, additional coordination beyond the daily planning period may be required. Cooperative teaching also occurs among bilingual teachers, who may trade groups for instruction in various content areas. Thus, the daily planning period is essential to track progress and make plans for adjusting the instructional program for each child. Teacher aides are frequently involved in these planning sessions.

Curricula, materials, and time involved.

Kindergarten -- Children are not screened for eligibility, and participation in the program is voluntary. During the first weeks of school, kindergarten teachers closely observe children in class and at play to determine the extent to which they rely on their native Spanish language for communication. Their ability to comprehend English is one basis for grouping and regrouping children for the remainder of the year.

In all project kindergartens, nearly two hours every day are devoted to language development and language arts in English and Spanish. For example, 45 minutes in the morning are used for English and Spanish language development during a "sharing experiences" period, discussion of the day, date, weather, and other calendar-related events, and writing on the blackboard stories the children dictate to the teacher. Another hour in the afternoon is reserved for aural-oral and reading readiness skills in English[1] and Spanish, using very detailed teacher guides developed especially for the program. In both languages, the patterned practices stress basic sentence patterns and illustrate changes in word forms and word order. Correct pronunciation and intonation are stressed during these drills. In contrast, during the times when free expression is encouraged, teachers do not correct the children or in any way evaluate their performance. The emphasis at these times is on creating excitement and adding fuel to the children's growing enthusiasm for communicating in English and in Spanish.

---

[1] English reading readiness is introduced as kindergarteners are able to learn these skills. For 80% of the children, English reading readiness skills are taught in kindergarten; for the other 20%, English reading readiness skills are introduced early in Grade 1.

Other periods of time are set aside for activities which lend themselves to further development of communication skills. For example, number concepts, story time, filmstrips, music rhythms, snack time, and cleanup are used to develop communication in English and Spanish. Even during independent work-play activities, the teacher and aide circulate to provide each child with at least a brief opportunity to discuss his activity in both languages. Thus, from the beginning, the program emphasizes communication skills heavily, and provides for their systematic development in Spanish and in English throughout the daily schedule of activities.

Commercial materials used in addition to the specially prepared Spanish and English aural-oral guides are reading readiness materials by Houghton-Mifflin and Harcourt Brace Jovanovich, Inc. Mathematics textbooks are by Addison-Wesley (Spanish edition) and Houghton-Mifflin.

For most kindergarteners, instruction is provided in Spanish for 90% of the day and in English for 10% of the day. Individual differences among children necessitate adjustments in language emphasis, and by the end of the kindergarten year, many children are receiving instruction in the two languages for equal amounts of time. Before the end of Grade 1 this is true for all children.

Grade 1 -- Each day about one hour is devoted to Spanish reading and language arts, and about two hours to English reading and language arts. The time an individual child spends on these activities fluctuates depending on his need. Other significant blocks of time are scheduled daily for Spanish social studies (emphasis on Spanish culture and heritage) and for mathematics--about three quarters of an hour each.

Phonetic analysis skills are taught first in Spanish because of the highly phonetic nature of the language. The teachers use detailed lessons and activities specified in the specially prepared Spanish reading guide. The guide, divided into several levels, also specifies charts the teacher must make to accompany the carefully sequenced lessons. In the early lessons on vowel sounds, the guide also presents dialogues the teacher uses to get the children to talk about the sound and to practice saying it and recognizing it in oral speech. Each lesson in the guide also suggests supplementary activities the teacher or aide can use to reinforce the lesson. After the children have learned the vowels and can recognize them as initial sounds or as sounds within words, they learn consonants in the same fashion, according to sequenced lessons in the guide. Lessons then progress to sounds of letter combinations and syllables. By the end of eight to ten weeks, every child will have completed the unit on phonetic analysis skills and will have begun to decode words. Other reading skills are also taught through lessons in the Spanish reading guide. Once children are ready to read in Spanish, the Laidlaw series is used.

English reading readiness skills for Grade 1 youngsters are taught through the Harcourt Brace Jovanovich, Inc. series. Phonetic analysis skills are introduced after the child has learned them through his Spanish reading lessons. Teacher guides accompanying the commercial English reading series provide instructions for conducting the daily lessons. There is a heavy phonics emphasis, with comprehension skills receiving heavier emphasis as decoding skills develop. Other commercial materials used for English reading are by Houghton-Mifflin (Listen and Do) and by Scott Foresman (Talking Alphabet). By the last six weeks of the school year, the top reading group in each first grade is reading second grade books in the Harcourt Brace Jovanovich, Inc. series.

In Grades 1 through 4, mathematics is taught using the Addison-Wesley series in English, but to help provide instruction in Spanish as well, the teacher also uses a companion teacher's edition of the series in Spanish. In Grades 1 through 4 the teacher-developed Spanish social studies guide is supplemented with lessons in English social studies from a textbook by Harcourt, Brace, and World.

During the time devoted to these subject areas, English and Spanish may be used interchangeably. There is no regimen for this, and teachers use various strategies; in general they switch back and forth between the languages depending on the difficulty the children are having with the concept or subject matter, or a teacher will devote one week to emphasizing initial instruction in Spanish with reinforcement of the concepts in English, then one week to the opposite order of presentation. Before the end of Grade 1 English and Spanish are being used equally.

Grades 2, 3, and 4 -- In Grades 2 and 3 Spanish reading continues to be taught according to detailed lessons in more advanced levels of the special Spanish reading guide, and the Laidlaw series continues to be used through Grade 4. The Harcourt Brace Jovanovich, Inc. series is used throughout for English reading. Although less time is devoted to reading and language arts in both languages than was true of Grade 1, these skills continue to be emphasized in Grades 2, 3 and 4. By the last six weeks of the year, children in the best reading group in each class are reading books on the next grade level.

Grouping strategies and individual attention, Grades 1-4 -- To increase the effectiveness of instruction in both languages in the elementary grades, children are grouped for activities in the various subject areas according to language dominance and achievement. There are usually three or four groups working separately in every classroom, one with a teacher, one with her aide, while the other one or two groups work on independent assignments made at the beginning of the day. Groups are small, usually seven or eight children. The groups rotate every 20 or 30 minutes so that each child receives a substantial portion of instruction in each subject from the teacher and reinforcement from the aide. In addition, children who need extra coaching in reading and language arts receive special, individual help from the teacher or the aide while the rest of the class is working on independent seat work, projects, or daily assignments. At least one parent is usually in the classroom to listen to children read aloud, to keep an eye on youngsters who are working independently, and to chat in Spanish with them about their activities and interests.

After-school study centers, Grades 3-4 -- In 1974-75, program participants in Grades 3 and 4 had access to further individual help from one of the project teachers for one hour after school at each of the participating schools. Up to 15 children from each project class of third and fourth graders were accepted for this hour of personalized assistance in problem areas in reading, language arts, and mathematics.

Facilities. No remodeling of existing classroom facilities was required to implement the program. In one project school classes are self-contained. In the second school, all classes at a grade level are conducted in an "open space" large enough for teachers to instruct their classes in different areas of the room. Regrouping and cooperative teaching (daily trading of groups and teachers for particular subject areas) is very convenient under this architectural plan. In the third project school, the building is circular, with "pie-shaped" classrooms having access to a central resource center.

Preservice/inservice training. A week of concentrated preservice training is required for teachers and teacher aides prior to the beginning of school each year. During the morning, all staff hear guest speakers from universities lecture on areas related to project philosophy, objectives, and instructional procedures. During the afternoon the speakers act as consultants during workshop sessions in which teachers apply principles and procedures presented during the lecture. In these sessions, teachers share ideas, develop materials, demonstrate techniques, and discuss problem areas; teachers are usually grouped by grade level for these specialized sessions. All project teachers and teacher aides are paid for participating in the preservice training week, and there has been 100% participation every year.

During the year one day of inservice training is required for teachers every month during school hours; substitute teachers cover project classes on these days so that participation of all teachers is possible. The inservice sessions deal with interests teachers indicate to the project administrator and Instructional Consultants and provide further opportunities for idea sharing and refinement of instructional procedures.

Every preservice and inservice session is critiqued by teachers and teacher aides on special rating forms. Their feedback is used to modify future training sessions and to plan additional programs of special interest.

Another feature of staff development is the use of the Flanders Interaction Analysis with new teachers. Videotapes are made of classroom lessons in September, January, and May. The project evaluator analyzes the tapes and confers with each teacher to review and discuss the findings. These conferences are private, and both negative and positive aspects of the tapes are reviewed. The teachers are guided to define their own concept of ideal teacher behavior and subsequently to modify their own behavior in the direction of that idea. Since the focus of the project is on language development, particular attention is paid to the percentage of student talk and of indirect teacher approach.

Parent-community involvement. Four to five parent meetings are held at each of the project schools every year, led by one of the Instructional Consultants. Topics for these meetings include parental involvement at home and school, teaching aides, the Toy Lending Library, and the goals and objectives for the project. Through these meetings and through contact with the teachers, the parents are encouraged to help with classroom activities. They help to reinforce academic work, supervise playground play, chaperone field trips, bring snacks, tell stories, make charts and other instructional aids, and talk with children informally in class. Parent involvement has increased dramatically since the program began--from nearly total noninvolvement to donating over 3300 hours of help in classroom-related activities in 1973-74. A Parent Advisory Committee (PAC) is elected each year and consists of 10 parents, the project administrator, and the Instructional Consultants. In 1974-75 there were 5 parents from one school, 3 from the second school, and 2 from the third, representing Grades K through 4. The PAC meets once a month to hear progress reports on the program, to discuss what the program should be accomplishing, to review grant applications, and to discuss ideas for modifications in the program. One idea that was incorporated in the 1975-76 grant application was preservice and inservice training for parent volunteers.

Cost. Budget information provided by the project administrator was based on Title VII allocations and expenditures for four years. For the first three years of program operation, this amounted to nearly 100% of the budget. A summary of total expenditures, enrollment, and associated per pupil costs for four years is given below.

	Enrollment	Total Expenditures*	Per Pupil Cost
Year 1	143	$65,460	$458
Year 2	290	$60,265	$208
Year 3	435	$86,025	$198
Year 4	519	$82,202	$158

*Excludes teacher salaries and salaries for teacher aides in Title I schools; only Title VII expenditures are shown.

In order to compare per pupil cost for the total educational program of children in the Title VII project with the comparable figure for children in the regular program, further calculations were made. For 1973-74, based on total funds allocated for the program by federal, state, and local sources (page 1 of this summary), the per pupil cost for the bilingual program was $622 for each of the 519 children served. A finance officer for the Corpus Christi School District indicated that the comparable per pupil allocation in 1973-74 for children not enrolled in the bilingual project was $541 per child in the elementary grades. Both of these per pupil cost figures include all allocations for the total educational program for these two groups of children, except amounts budgeted for "maintenance of operations."

The drop in per pupil costs over the first three years, when Title VII funds covered almost all program costs but instructional salaries, reflects the growing emphasis on carefully reviewing expenditures for travel, materials, and equipment and on sharing available resources. A major expense has been paying teachers to develop the Spanish and English aural-oral kindergarten guides, the Spanish social studies guide, and the Spanish reading guide, Levels I through VII. Another major expense each year has been paying project teachers and aides to attend the preservice training week, and paying substitute teachers to cover classes for project teachers on inservice days during the year. Economies have resulted from having parents repair instructional materials and aids and make teaching aids.

EVIDENCE OF EFFECTIVENESS:

Summary

The evaluation data summarized in this section will be based upon the evaluation reports for the 1972-73 and 1973-74 academic years. The data strongly suggest that the bilingual program in Corpus Christi results in superior achievement in both languages by project pupils in comparison with controls. It is useful to summarize the results by grade level for the two years.

1972-73

- Kindergarten. Project pupils showed significantly higher adjusted posttest scores on the Inter-American General Ability Spanish subtests than did controls. There were no significant differences on the two Inter-American English subtests nor on three subtests of the Stanford Early School Achievement Tests.
- First Grade. Project pupils showed significantly higher adjusted posttest scores on both subtests of both the Spanish and English versions of the Inter-American Reading Tests, and also achieved significantly higher grade equivalents on all three subtests of the SRA Achievement Tests.
- Second Grade. Project pupils showed significantly higher adjusted posttest scores on both subtests of both versions of the Inter-American Reading Tests and on two of the three subtests of the SRA Achievement Tests.
- Furthermore, with but one subtest exception, the comprehension part of the Spanish version of the Inter-American Reading Test, length of enrollment in the program was shown to be significantly related to test performance.

1973-74

- Kindergarten. Project pupils showed significant gains in comparison to controls on the Spanish version of the Inter-American General Ability Test, but were equivalent to controls on the Stanford Early School Achievement Test. Objectives were approached but not reached at the kindergarten level.
- First Grade. First graders in the project outperformed controls by approximately 0.5 grade equivalents on the SRA Achievement Tests (significance levels are not reported), but did significantly outgain controls on both versions of the Inter-American Reading Tests.
- Second Grade. Project pupils outperformed the controls with respect to average grade equivalents on the SRA Achievement Series (though the mean differences were not statistically significant). Project pupils significantly outgained controls on the Spanish version of the Inter-American Reading Tests. The reverse was true on the English version of the Inter-American Reading Test (controls significantly outgained project pupils).

- Third Grade. While all differences on the three subtests of the SRA Achievement Tests favored the project group, there were none significant. However, project pupils outgained the controls on both versions of the Inter-American Reading Tests.

Background

The Title VII Bilingual Education Program in Corpus Christi, Texas is in the fifth year of operation. During the 1973-74 academic year, the last full year for which evaluation data is available, 519 pupils were enrolled in the project in grades K through 3 in each of three schools (Crockett, Evans, and Travis). Table I shows the growth of the Title VII project during its first four years; one grade level was added each year.

Table 1

Number of Students Enrolled at Each Grade Level Each Academic Year in the Corpus Christi Bilingual Education Program

	Grade Level				Total
Year	K	1	2	3	Enrollment
1970-71	143	----	----	----	143
1971-72	184	106	----	----	290
1972-73	173	155	107	----	435
1973-74	154	148	124	93	519

Evaluation

The Title VII evaluator in 1972-73 and in 1973-74 was Ilene Cordray. She is also a staff member of the Corpus Christi School District Department of Planning and Evaluation, directed by Dr. Ronald Howard in 1972-73, and by Dr. Kenneth Kyle in 1973-74.

Pupil Characteristics

In 1973-74, the total C.C.I.S.D. student enrollment was slightly under 44,000 with approximately 32,500 Mexican-American pupils (54%), 2,400 black pupils (5%) and 17,700 Anglo pupils (41%). The 519 Title VII project pupils were 92% Mexican-American, 6% Black and 2% Anglo. Two of the three project schools show similar ethnic balances with the exception of the Evans school with an enrollment of 75% Mexican-American, 24% Black and 1% Anglo.

Data reported for the 1972-73 academic year indicated low average monthly family income levels of $356 for project enrollees with 81% of the pupils' family incomes falling below the poverty level. Average family size included 4.25 children based on parents' reports; 323 pupils (74% of the 435 enrollees) were native Spanish speakers. Average class enrollments in September 1972 were 27.75; withdrawals during the 1972-73 year averaged 4.81 per class (17 per cent); late entries per class averaged 4.00 (14%). Absences during the year were high, averaging approximately 353 per class or 12.7 per pupil.

Three elementary schools selected by the Title VII staff served as control schools. One class at each grade level was selected within each of the three control schools by the principal in charge of each. Classes forming the control group were approximately equal to the project group in terms of socioeconomic level. Each control school was also the one closest geographically to a project school.

Data reported for 1972-73 also indicated that kindergarten teachers in the control schools averaged 20 years' teaching experience, while the six kindergarten teachers in the project schools were all in their first or second year of teaching. Class size for control kindergartens averaged 23 in comparison with an average size of 29 for project kindergarten classes. Otherwise, at grades 1 and 2, project and control classes were equal in terms of teachers' experience and class size.

Until the current 1974-75 academic year, classes in the control schools were taught monolingually without regard for the language dominance of the children. With the implementation of Texas bilingual program legislation in 1974-75, pupils in the control classes may also be partially instructed in Spanish. The project evaluator reports that in the control schools, heavy emphasis was placed on the instruction of reading and mathematics just as in the Title VII bilingual education program.

Comparison Design

The basic comparison design in 1972-73 and 1973-74 involved administration of pretests and posttests to both project and control groups within each grade. The evaluators point out that the two groups may not show sampling equivalence although the attempt was made to balance both groups as closely as possible in terms of ethnic mix, maturation, and the test instruments administered. Matching was made at the school level. In this situation, if pretest equivalence is observed for the controls and project pupils, then significant mean posttest scores, either raw or adjusted for pretest group differences, or significant mean gain scores favoring the project pupils indicate program effect, other things being equal. Data will be reported for the 1972-73 and 1973-74 academic years.

Measures

For the 1972-73 and 1973-74 academic years, the following tests were administered at the grade levels indicated.

Test	Grades administered in:	
	1972-73	1973-74
Stanford Early School Achievement Test	K	K
Inter-American General Ability (Spanish)	K	K
Inter-American General Ability (English)	K	K
SRA Achievement Test (Primary I)	1	1
SRA Achievement Test (Primary II)	2	2,3
Inter-American Reading (Spanish Level I)	1,2	1,2
Inter-American Reading (Spanish Level II)	-	3
Inter-American Reading (English Level I)	1,2	1,2
Inter-American Reading (English Level II)	-	3

Results: 1972-73

Kindergarten. Table 2 shows means on the Stanford Early School Achievement Test (May administration) for the kindergarten project and control pupils enrolled during the 1972-73 academic year. Table 2 also shows mean pretest, posttest, and posttest-adjusted-for-pretest group mean differences on the verbal-numerical and non-verbal parts of both the Spanish and English language versions of the Inter-American General Ability Test.

In kindergarten, both the control and project group showed equivalent means on the three subtests of the Stanford Early School Achievement Test and on both tests of the English versions of the Inter-American Test of General Ability. On the Spanish version of the Inter-American (Pruebas de Habilidad General), the adjusted mean posttest differences significantly favored the project pupils on both subtests. In both English and Spanish language development the project pupils scored above the 64th percentile on both language versions of the Inter-American Test, while the controls scored significantly lower (at the 46th percentile) on the Spanish version.

Table 2

Means on the Stanford Early School Achievement Test and on the Inter-American General Ability Test (Spanish and English Versions)[1] for Bilingual Project and Control Group Students

Grade (Language)	Group	N	Stanford Early School Achievement Test		
			Mathematics	Letters, Sounds	Aural Comprehension
Kindergarten (English)	Project	167	17.4	17.6	17.0
	Control	67	18.6	18.0	18.0

			Inter-American Pruebas de Habilidad General					
			Verbal-Numerical			Non-Verbal		
			Pre	Post	Post(adj)	Pre	Post	Post(adj)
Kindergarten (Spanish)	Project	143	9.6	12.9	13.0**	11.5	15.2	15.2*
	Control	59	10.3	12.1	11.8	12.1	14.5	14.3

			Inter-American Test of General Ability (English)					
			Verbal-Numerical			Non-Verbal		
			Pre	Post	Post(adj)	Pre	Post	Post(adj)
Kindergarten (English)	Project	143	10.4	14.4	14.3	12.0	15.1	15.1
	Control	59	10.2	14.0	14.1	11.0	15.3	15.3

* Mean differences between Project and Control group at .05 level

** Mean differences between Project and Control group at .01 level

---

[1] Raw score means are shown for the Inter-American General Ability Test.

First Grade. First grade summaries are shown in Table 3 which reports grade equivalent means for project and control groups on the three subtests of the SRA Achievement Test and pretest, posttest, and gain score means on both subtests of both language versions of the Inter-American Reading Tests.

Table 3

Means on the SRA Achievement Test and on the Inter-American Reading Tests (Spanish and English Versions)[1] for Bilingual Project and Control Group Students

		SRA Achievement Test (Primary I) GE's		
Group	N	Reading	Language Arts	Mathematics
Project	148	2.2	2.1	2.6
Control	75	1.8**	1.7**	1.8**

		Inter-American Pruebas de Lectura (Nivel I)					
		Vocabulary			Comprehension		
		Pre	Post	Post(adj.)	Pre	Post	Post(adj.)
Project	126	9.3	20.4	20.4	9.5	16.2	16.2
Control	65	7.7	9.4	9.2**	8.3	8.7	8.8**

		Inter-American Test of Reading (Level I)					
		Vocabulary			Comprehension		
		Pre	Post	Post(adj.)	Pre	Post	Post(adj.)
Project	126	10.9	23.9	23.4	9.9	19.6	19.4
Control	65	8.5	18.0	18.9**	7.8	13.6	14.1**

** Mean differences between Project and Control group at .01 level

Both the project and control groups performed well at the end-of-year administration of the SRA Achievement Tests (1.8 = appropriate grade level), but the project group scored significantly higher (.01 level) in mean grade equivalents than did the controls on all three subtests. On both language versions of the Inter-American Reading tests, the adjusted posttest means of the project pupils were significantly higher (.01 level) than the means of the control group on both the vocabulary and comprehension subtests.

---

[1] Raw score means are shown for the Inter-American Reading Tests.

Second Grade. Second grade summaries are shown in Table 4 for the same tests as those appearing in Table 3.

Table 4

Means on the SRA Achievement Test and on the Inter-American Reading Tests (Spanish and English Versions)[1] for Bilingual Project and Control Group Students

Grade (Language)	Group	N	SRA Achievement Test (Primary II) GE's		
			Reading	Language Arts	Mathematics
Second (English)	Project	106	2.9	2.8	2.7
	Control	68	2.6	2.5*	2.4*

			Inter-American Pruebas de Lectura (Nivel I)					
			Vocabulary			Comprehension		
			Pre	Post	Post(adj.)	Pre	Post	Post(adj.)
Second (Spanish)	Project	92	14.6	26.4	25.8	10.3	21.1	21.4
	Control	68	12.3	16.4	17.2**	11.4	14.6	14.2**

			Inter-American Test of Reading (Level I)					
			Vocabulary			Comprehension		
			Pre	Post	Post(adj.)	Pre	Post	Post(adj.)
Second (English)	Project	92	22.4	34.0	33.7	17.7	32.8	32.4
	Control	68	21.3	30.7	30.9**	18.3	27.7	27.5**

* Mean differences between Project and Control group at .05 level

** Mean differences between Project and Control group at .01 level

The project students maintained grade level (2.8) at the May testing of the SRA Achievement tests and significantly (.05 level) out-scored the controls on the SRA language arts and mathematics subtests. On both the English and Spanish versions of the Inter-American Reading tests, the project group's adjusted posttest means were significantly higher (.01 level) than the control's on both the vocabulary and comprehension tests.

---

[1] Raw score means are shown for the Inter-American Reading Tests.

Language Development. Project pupils at the end of the first grade were classified into those who had been in the program for two years and those who had been in the program only for the first grade. On all three subtests of the SRA Achievement Test--Reading, Language Arts, and Mathematics--the pupils in the first grade who had also been enrolled in the bilingual program in kindergarten scored significantly higher than pupils enrolled for only the first grade. The same pattern of significant differences between SRA Achievement Test means for second graders enrolled three, two, and one year in the program was observed. These data are shown in Table 5.

Table 5

Mean SRA Achievement Test Scores for First and Second Graders Classified by Length of Enrollment

		SRA Achievement Test		
Group	N	Reading	Language Arts	Mathematics
First Graders in Program for:				
2 years	100	2.3**	2.3**	2.7*
1 year	23	1.9	1.8	2.4
Second Graders in Program for:				
3 years	42	3.1**	3.3**	2.8**
2 years	21	2.8	2.5	2.5
1 year	27	2.6	2.4	2.5

* Mean differences between groups at .05 level

** Mean differences between groups at .01 level

Mean pretest and posttest raw scores on both the English and Spanish versions of the vocabulary and comprehension subtests of the Inter-American Reading Tests were also calculated and are reported in Table 6 for first graders enrolled in the project for one or two years, and for second graders enrolled one, two, or three years. In every instance, with the exception of the Comprehension posttest scores for first graders on the Inter-American Spanish version, length of program enrollment was significantly related to mean test performance for both grade groups.

Table 6

Raw Score Means on the Inter-American Reading Test for
First or Second Graders Classified by Length of Enrollment

		Inter-American Reading Test							
		Spanish				English			
		Vocabulary		Comprehension		Vocabulary		Comprehension	
Group	N	Pre	Post	Pre	Post	Pre	Post	Pre	Post
First Graders Enrolled:									
2 years	102	9.8	21.6	9.6	16.7	11.2	25.0	10.2	21.0
1 year	22	7.3	14.6 **	9.2	14.2 n.s.	9.6	19.4 **	8.9	13.7 **
Second Graders Enrolled:									
3 years	43	16.9	29.7	12.6	24.9	26.1	36.0	20.3	35.3
2 years	21	16.5	26.0 **	10.9	19.3 **	21.2	32.9 **	17.3	30.0 **
1 year	28	9.3	21.2	6.0	16.1	17.9	31.7	14.1	29.5

** Mean differences between groups at .01 level

Results: 1973-74

Comparisons between project and control groups in 1973-74 were made on the same tests used in 1972-73 to measure English and Spanish language abilities. However, posttest covariance adjustments were not calculated, so that on the Inter-American Series tests in both languages, pretest, posttest, and gain means are reported, with program effectiveness demonstrated by significant mean gain differences favoring project groups.

Kindergarten. Certain criteria were set as indicative of the achievement of program objectives for the project group. On the Stanford Early School Achievement Test, kindergarten pupils enrolled in project classes in 1973-74 were expected to achieve at an 80/40 (80% of the pupils will score above the 40th percentile) criterion and a 40/70 (40% of the pupils will score above the 70th percentile) criterion. On the Spanish version of the Inter-American General Ability Test the mean gain of the project group was expected to show a significantly (.05 level) higher mean gain than the controls with 90% of the pupils scoring above the 20th percentile and 60% of the pupils scoring above the 50th percentile.

For the Stanford Early Achievement Test total score, 83% of the project students scored above the 40th percentile and 27% of the project students scored above the 70th percentile. Mean differences on the three subtests did not significantly favor the project students in any of the three subtests.

For the Spanish version of the Inter-American Series posttest total score, 85% of the project pupils scored above the 20th percentile and 70% above the 50th percentile, indicating near attainment of the achievement expectancies.

Table 7 shows pretest, posttest, and gain means on both versions of the Inter-American General Ability Test for the project and control students.

Table 7

Raw Score Pretest, Posttest, and Gain Means for Kindergarten Pupils

		General Ability Inter-American Total Spanish			General Ability Inter-American Total English		
Group	N	Pre	Post	Gain	Pre	Post	Gain
Project	99	22.3	28.8	6.5	22.0	30.4	8.4
Control	48	24.2	26.2	2.1^{*}	22.0	29.3	$7.3^{n.s.}$

* Mean differences between Project and Control group at .05 level

In summary, objectives were partially achieved at the kindergarten level, with pupils in the project significantly outgaining controls on the Spanish version of the Inter-American Series.

First Grade. Objectives for the first grade were (a) that first grade project pupils would significantly outgain control pupils on both versions of the Inter-American Reading Test, and (b) that project pupils would score 0.6 grade equivalents higher than controls on the SRA Achievement Mathematics and Reading subtests. Summary test information is presented in Tables 8 and 9.

Table 8

SRA Achievement Test Means for First Graders

		SRA Achievement Tests		
Group	N	Reading	Language Arts	Mathematics
Project	111	2.3	2.2	2.4
Control	74	1.8	1.7	1.8

With respect to the second objective, grade equivalent differences favored the project group on the Mathematics subtest at the expected level (0.6), but not on the Reading subtest (0.5) of the SRA Achievement Test. Significance levels are not reported for these mean differences between project and control pupils in the first grade, but it should be noted that project pupils averaged considerably above the grade level of 1.8.

Table 9

Inter-American Reading Test Raw Score Means for First Graders

		Inter-American Spanish			Inter-American English		
Group	N	Pre	Post	Gain	Pre	Post	Gain
Project	125	19.6	49.5	29.9	22.2	52.9	30.7
Control	46	14.4	21.3	6.9**	14.2	34.4	20.2**

** Mean differences between Project and Control group at .01 level

With respect to the total score on both versions of the Inter-American Tests (English and Spanish) the project pupils showed significantly higher gain scores than did the controls in the first grade.

Second Grade. Objectives for the second grade were (a) project students would average 0.6 grade equivalents higher than controls on the Mathematics and Reading subtests of the SRA Achievement Test and (b) project students would outgain controls on both versions of the Inter-American Reading Tests. Second grade project students outperformed the control students on the SRA Achievement Tests, although the expected differences of 0.6 grade equivalents were not observed, as shown in Table 10.

Table 10

SRA Achievement Means for Second Graders

		SRA Achievement Tests		
Group	N	Reading	Language Arts	Mathematics
Project	130	3.1	3.1	2.9
Control	83	2.7	2.6	2.5

Table 11 shows mean pretest, posttest, and gain scores for both the project and control students on both versions of the Inter-American Reading Tests. Results were mixed with project pupils outgaining controls on the Spanish version and controls outgaining project pupils on the English version.

Table 11

Inter-American Reading Test Raw Score Means for Second Graders

		Inter-American Spanish			Inter-American English		
Group	N	Pre	Post	Gain	Pre	Post	Gain
Project	97	33.5	54.0	20.5	44.6	65.6	21.0
Control	57	20.4	35.9	15.5**	34.3	60.8	26.5**

** Mean differences between Project and Control group at .01 level

Third Grade. Third grade objectives were identical to those for the second grade: (a) 0.6 mean grade equivalent superiority for project pupils on the Mathematics and Reading subtests of the SRA Achievement Tests and (b) mean gain score superiority for project pupils on the Inter-American Reading Tests. Results are presented in Table 12 and Table 13.

Table 12

SRA Achievement Test Means for Third Graders

		SRA Achievement Tests		
Group	N	Reading	Language Arts	Mathematics
Project	87	3.5	3.6	3.6
Control	89	3.1	3.5	3.5

Objectives with respect to the SRA Reading and Mathematics subtests were not met and grade equivalent levels for either group do not reach grade expectancy. However in reading, the project pupils outperform the controls by 0.4 grade equivalents on the average.

Table 13

Inter-American Reading Test Raw-Score Means for Third Graders

		Inter-American Spanish			Inter-American English		
Group	N	Pre	Post	Gain	Pre	Post	Gain
Project	85	26.7	56.5	29.8	52.5	72.7	20.2
Control	63	23.8	33.5	9.7**	48.2	59.1	10.9**

** Mean differences between Project and Control group at .01 level

In the third grade, project pupils significantly outgained the controls on both versions of the Inter-American Series.

PROGRAM AREA:

Bilingual Education

PROJECT TITLE:

Bilingual Education Program (Title VII, PL 89-10)

LOCATION:

Houston, Texas

SOURCES AND LEVEL OF FUNDING:

The following figures pertain to the fifth year of program operation:

Title VII Funds	$246,960
Local Funds	182,344
State Funds	296,705
Total Year 5	$726,009

PROGRAM START DATE:

1969

BRIEF DESCRIPTION OF PROJECT:

Goals and objectives. The program is designed to serve Spanish speaking students by developing culturally and linguistically appropriate curriculum, by providing relevant training for teachers and aides, by developing parent and community involvement in the educational process, and by generally effecting a change in attitude towards bilingual education. The intent of this program is to improve student self-concept by raising the status of the students' language and culture and thereby increasing student achievement in English, Spanish, and other content areas.

Context. The Houston school district, which is 41% Black, 40% Anglo, and 18% Mexican-American, experiences a high attrition and mobility rate. The program schools are located in the eastern and northeastern parts of the city and have a student enrollment which is 53% Mexican-American, 42% Black, and 5% Anglo, Asian, and Indian.

Program description.

Grade level(s), years of operation, size--During its fifth year of operation, 1973-74, the program served 1550 K-12 students in eight elementary schools, one junior high, and one high school.

Staffing--The program staff required for this number of participants consists of the following: Administrators/supervisors (5, full time), Specialist (1, parttime), Evaluator (1, parttime), teachers (48, full time), teacher aides (23, parttime), and clerical (3, full time). All administrators, teachers, and aides in the program are bilingual.

Curricula, material, time involved--(This section applies to the elementary level program.) Students begin learning to read in Spanish with the use of Mis Primeras Letras, a reading readiness book on letters and letter sounds in Spanish. Spanish Reading Activities, a workbook developed by teachers in the district, is used in conjunction with the reading text. After students master Spanish decoding skills using Mis Primeras Letras and Spanish Reading Activities, they begin reading in the Laidlaw basal reading series in Spanish. As students improve their reading skills in Spanish, they progress through the Laidlaw basal readers.

Spanish dominant students are also instructed in oral English language development, and English reading using Basic English Language Patterns for Spanish Speaking Students and the Harcourt, Brace, and Javonovich basal reading series. English dominant students are instructed in oral Spanish, Spanish reading (for those who are ready), and English reading.

Each bilingual classroom has a teacher and a half-time aide who assists with instruction. The teacher plans the daily program, and the aide works closely with her to see that it is carried out.

Instruction includes a block of time devoted to Spanish reading and language arts. During the remainder of the day the instruction is in English, for English dominant and bilingual students. Students who are monolingual Spanish receive additional instruction in Spanish after the lessons are presented in English. The bilingual teachers have available to them the flexibility of explaining, emphasizing, or reinforcing any skill in either language to ensure that each pupil becomes involved and senses a feeling of achievement and security.

At the secondary level, the program includes ESL for monolingual Spanish students and several bilingual courses for students who have already attained some degree of bilinguality. The monolingual Spanish students are identified by the homeroom teachers and tested on basic reading skills. Those students who qualify are enrolled in ESL, plus regular English and math, and their choice of elective courses. Students in their second year of ESL also take Science and History (in English).

Students who do not qualify for ESL may enroll in any of the several courses which are taught bilingually. These courses are open to all students.

Bilingual Courses

Junior High	High School
Reading, Writing, Spelling	Algebra
Texas History	Geometry
Speech	Spanish Business Communication
Drama	Mexican Folkloric Dance
Spanish for the Spanish Speaker	

Facilities--The program can operate in a regular classroom without any special modification.

Parental involvement--Close contact with parents has been a key feature of the program over the five years of its development. Each year teachers hold quarterly meetings with the parents at the school to explain the bilingual program. Teachers also make yearly home visits to parents to talk about how their child is doing and to obtain the parents' reactions to the program on a district questionnaire.

Brochures entitled "Parental Involvement in the Bilingual Education of Houston Independent School District" have been developed by the district to guide teachers in getting parents involved in school-related activities. The teachers have recruited parents to help on various classroom projects, and some parents have tutored bilingual students under the guidance of the teachers.

Advisory Board Meetings--During the first few years of the program's development, the district Title VII Advisory Board played an important role in increasing community involvement in the program. For instance, in order to increase participation of Spanish speaking parents, meetings were conducted in English and Spanish. At the beginning these meetings were used to inform parents of developments in the program. Later, parents became more actively involved in decision making.

The Advisory Board, which meets quarterly, consists of the principal, one bilingual teacher, one parent, and one community leader from each of the ten participating schools. These members are organized into study teams in six areas: Instruction, Curriculum, Community Relations, Inservice, Dissemination, and Evaluation. Each team reviews the needs and/or problems related to its area and makes recommendations to the total board.

Preservice/inservice training--Over the first five years of the program, teachers gained valuable experience in many areas of bilingual education. They participated in the development of new materials, established their own styles of dealing with the task of teaching in two languages, and discovered how to make the best use of the help an aide could provide in the classroom. The frequent inservice training sessions which were provided by the district for the teachers supported the teachers in their efforts.

This inservice training for the teachers and aides in the bilingual program was provided at least once a month during the school year. Training consisted of 2-hour sessions after school for which participating teachers received a $10 stipend. During the 1973-74 school year ten sessions were held, covering the following topics: ESL, cultural awareness and history, evaluation, and methodology of language instruction. Bilingual staff development activities also included visits to other schools, displays of locally developed instructional materials, and mini-university courses offered by the district.

Cost--During Year 1 the program received funds totaling $234,850; this figure has increased each year by about $100,000 as more participants are served. (Program enrollment has increased from 650 to 1550 participants over the five years.) The approximate per-pupil cost for the program is less than $500 per pupil, based on a five-year average. Major startup costs were for equipment and furniture ($20,000 in Years 1 and 2 combined), instructional materials ($10,000 in Year 1), inservice stipends ($14,500 in Years 1 and 2 combined), and consultants who helped with all phases of curriculum development, inservice training, and evaluation ($6,000 in Year 1). After Year 1 the amount spent for consultants dropped to $2,000 and continued to drop as local resource people gained experience in the program and were able to take over a consulting role. As the need for inservice training tapered off, the amount spent on stipends was reduced. Teachers are paid to write curriculum materials to supplement commercial materials, but this activity has become less critical as better bilingual materials become available commercially.

The Houston Independent School District substantially increased its support of the program during Years 3, 4, and 5 by picking up all program expenses at Kindergarten in Year 3, Kindergarten and Grade 1 in Year 4, and Kindergarten through Grade 2 in Year 5. Program staff feel that by assuming fiscal responsibility for the total program, a grade level at a time, year by year, the district funds have contributed to the orderly, stable growth of the program. This approach is believed to be preferable to one in which a local district picks up specific budget items, without taking into account sufficiently the impact of this action on the total program.

EVIDENCE OF EFFECTIVENESS:

Evaluation. The Bilingual Education Project of the Houston Independent School District is served part time by the project evaluator, Ms. Piedad Cortez, as well as a part-time research consultant who is responsible for the basic comparison design. The main objectives of the evaluation each year were to:

- Establish comparability of program participants and control groups to validate comparative cognitive gains.
- Compare the cognitive gains in general English ability of program participants and control groups.
- Assess gains in the Spanish ability of project participants.
- When possible, compare the mathematics ability of program participants and control groups.
- Compare mean posttest general English ability scores for control groups and for program participants categorized by years of enrollment in the program.

Year and grade samples. Elementary level test data for three years are summarized: 1971-72, 1972-73, and 1973-74. These are respectively the third, fourth, and fifth years of program operation. Data for grades K-3 are reported for 1971-72, grades K-4 for 1972-73, and grades K-5 for 1973-74. In 1973-74 however, the grade 5 comparisons are based upon only 36 pupils in the combined participant and control groups; this number is too small to allow strong inferences to be made. Consequently, the statistical summaries presented in Table 2 and Table 3 below are those for grades K-3 in 1971-72 and grades K-4 in 1972-73 and 1973-74.

Measures. Measures reported are the general English ability and mathematics scores of the Inter-American Series Tests administered in both Spanish and English to the program participants and in English to the control groups. Reliabilities are unreported. This test is quite generally in use in bilingual education programs across the nation. Raw scores are reported. There are no standardized measures, grade equivalent scores, or U.S. norms available.

Comparison procedures.

- Pretests were administered in October and posttests in May of each academic year.

- Repeated measures analyses of variance were calculated, yielding tests of group effects (experimental versus control) in the between-subjects breakdown and main time effects (pre- versus posttests) as well as Time X Group interaction effects in the within-subjects breakdown. The latter one degree of freedom contrast is the one of interest--its significance indicates the rejection of the null hypothesis that experimental and control mean gains are equal. Mean gains for both groups are reported and indicate in all cases that the estimated gains for the experimental group exceeded that for the control group within each grade in each year reported; nearly all gain contrasts were significant.

- Inferential statistics were F-tests.

- Participant schools and control schools numbered seven and two, respectively, in 1971-72 and the same in 1972-73; there were eight experimental and two control schools in 1973-74. Ethnic distributions of total enrollees in all project schools for the three years reported were approximately balanced between Black and Spanish-American which resulted from "pairing" of schools to observe integration; this "balancing" in Houston was among the predominantly Spanish-American schools.

- Control groups are reported as selected on the basis of similarity to the experimentals in language, socio-economic level, and academic achievement. However, three factors should be considered when interpreting the results. First, individuals were removed from the control groups in order to match the pretest score distribution in each grade to the pretest score distribution for the experimental group. Second, attrition for the bilingual participants in the Houston Independent School District is high. Longitudinal attrition figures are presented in Table 1 showing the numbers of continuing program participants over the first four years of the bilingual program. For example, of the 290 kindergarten pupils enrolled in the bilingual program in 1969-70, 153 continued in the program for two years, 99 for three years, and 75 for four years. Because of attrition, the evaluator felt compelled to require significance at the .01 level. (These figures are taken from the application for continuation for the fifth year. Numbers are not extended or reported in the 1973-74 fifth year program evaluation report.)

Table 1

Bilingual Program Enrollment by Year and by Grade
Showing Second, Third, and Fourth Year Continuing Pupils

Grade	1969-70	1970-71	1971-72	1972-73
K	290	175	153	184
1	124	307(153[a])	163(111[a])	172(125[a])
2	24	121(62[a])	259(94[a],99[b])	183(62[a],101[b])
3		24(8[a])	100(36[a],33[b])	256(72[a],62[b],75[c])
4			12(7[a])	127(38[a],30[b],27[c])
5				29(9[a],7[b])

a--Second Year Participants; b--Third Year Participants; c--Fourth Year Participants

Third, in 1972-73 a significant number of pupils enrolled in an ESL Project 21 were discovered to be in the control group and were then consequently eliminated from the analysis. (This in effect removed the Spanish-dominant segment of the control group.)

- Tests were administered by elementary coordinators with the assistance of classroom teachers.

- Pretest and posttest results are reported for both program participants and for control groups with complete and non-zero scores.

Results. Table 2 shows the mean pre- and posttest Inter-American Series Test scores for the experimental and control groups in each grade for each year. General English ability and mathematics scores are reported. With the exception of grades 1 and 3 in the 1973-74 school year, all contrasts significantly favor the bilingual participants. Thus the evidence strongly supports the conclusion that the effects of the bilingual program are positive across grades for the three-year period reported. The mean gain differences favoring the program participants tend also to be large in relation to the estimated standard deviations of the gains (estimated as the square root of the within-subject variability across groups).

Table 3 reports pre-post and gain scores on the Inter-American Series General Ability Test, Spanish Version, for the program participants. No comparisons with controls are made. The data indicate that gains from October to May are large across grades and years.

Tables 2 and 3 are shown on the following pages.

Mean Inter-American General Ability Test scores for grade groups within each year were categorized either as control or by length of enrollment in the bilingual program. In most cases, the groups differed significantly although there is no apparent positive correlation between length of enrollment and mean score. Post-hoc contrasts were not calculated.

Summary. On the basis of the data reported for three consecutive years, the bilingual education program in Houston results in significantly higher seven-month gains in the English version for program participants in comparison to that of controls.

The design applied to summarize the data is simple, to the point, and is one that was planned as part of the management monitoring process. It is a summative procedure that could well be copied by many bilingual programs that do not stress the summative evaluation question with either clarity or understanding, and consequently do not permit any comparative assessment to be made.

Table 2

Experimental-Control Comparisons
on the Inter-American Series English Version

Grade	Group	Inter-American Series General English Ability N	Pre	Post	Gain	SD	F	P	Inter-American Series Mathematics N	Pre	Post	Gain	SD	F	P
1971-72															
K	Experimental	93	43.6	70.4	26.8	3.5	144.9	.001	-	-	-	-	-	-	-
	Control	47	46.2	58.3	12.1										
1	Experimental	119	49.4	62.8	13.1	5.5	6.4	.05	-	-	-	-	-	-	-
	Control	100	50.4	61.1	10.7										
2	Experimental	205	47.3	61.6	14.3	6.6	26.4	.001	-	-	-	-	-	-	-
	Control	93	49.1	57.4	8.3										
3	Experimental	79	59.2	74.8	15.6	6.1	31.5	.001	-	-	-	-	-	-	-
	Control	80	60.9	68.8	7.9										
1972-73															
K	Experimental	159	34.6	64.2	29.6	11.8	25.4	.001	-	-	-	-	-	-	-
	Control	38	43.1	57.5	14.4										
1	Experimental	146	44.6	61.8	17.2	11.0	20.5	.001	149	8.3	12.0	3.7	2.7	10.0	.001
	Control	60	49.4	55.8	6.4				59	9.2	11.1	1.9			
2	Experimental	161	45.2	61.8	16.6	8.1	44.5	.001	161	8.6	15.8	7.2	3.4	48.3	.001
	Control	53	49.6	53.9	4.3				53	11.1	13.0	1.9			
3	Experimental	218	56.1	67.2	11.1	7.6	8.9	.01	218	13.8	19.4	5.6	3.3	15.2	.001
	Control	83	58.4	65.3	6.9				83	15.4	18.4	3.0			
4	Experimental	98	62.8	84.5	21.7	14.7	27.5	.001	98	17.6	28.4	10.8	5.2	96.7	.001
	Control	88	63.6	69.3	5.7				88	19.0	19.0	0.0			
1973-74															
K	Experimental	104	46.5	64.4	17.9	7.1	59.0	.001	-	-	-	-	-	-	-
	Control	82	49.2	59.0	9.8										
1	Experimental	145	51.3	62.2	10.9	6.1	1.6	n.s.	145	9.0	12.3	3.3	2.7	0.3	n.s.
	Control	45	50.0	59.0	9.0				45	8.7	11.6	3.1			
2	Experimental	155	49.4	63.6	14.2	6.4	11.7	.05	155	10.3	16.7	6.4	3.3	13.2	.001
	Control	53	48.3	57.2	8.9				53	10.5	14.4	3.9			
3	Experimental	146	58.4	69.0	10.6	5.3	3.0	n.s.	146	14.7	21.7	7.0	9.9	2.1	n.s.
	Control	62	56.2	64.8	8.6				62	13.8	17.7	3.9			
4	Experimental	151	59.9	79.0	19.1	10.9	20.7	.001	150	17.0	24.4	7.4	5.0	16.7	.001
	Control	58	63.0	71.2	7.8				58	17.3	20.2	2.9			

Table 3

Pre- Post and Gain Scores for Experimental Groups on the Inter-American Test Spanish Version

Grade	Group	N	Pre	Post	Gain	SD	F	P
1971-72								
K	Experimental	111	42.3	68.9	26.6	8.1	595.3	.001
1	Experimental	130	47.7	59.0	11.3	6.5	194.1	.001
2	Experimental	205	44.5	58.5	14.0	6.8	434.8	.001
3	Experimental	79	60.7	76.1	15.4	6.6	216.5	.001
1972-73								
K	Experimental	159	36.2	61.0	24.8	13.0	287.7	.001
1	Experimental	146	47.6	58.1	10.5	8.7	104.5	.001
2	Experimental	161	45.9	58.7	12.8	7.2	241.6	.001
3	Experimental	218	47.6	61.8	14.2	11.2	172.6	.001
4	Experimental	98	57.1	77.3	20.2	13.2	115.1	.001
1973-74								
K	Experimental	127	46.2	65.7	19.5	6.2	605.9	.001
1	Experimental	168	43.2	58.0	14.8	9.7	196.6	.001
2	Experimental	172	46.1	62.8	16.7	13.9	12.9	.001
3	Experimental	163	54.7	69.5	14.8	12.8	11.7	.001
4	Experimental	162	54.1	72.9	18.8	14.2	137.4	.001

PROGRAM AREA:

Bilingual Education (Spanish/English)

PROJECT TITLE:

Kingsville Bilingual Education Program

LOCATION:

Kingsville, Texas

SOURCES AND LEVEL OF FUNDING:

The following figures pertain to the fourth year of program operation.

Title VII Funds	$ 67,727
Local Funds	145,000
State Funds	990
Total	$213,717

PROGRAM START DATE:

1970

BRIEF DESCRIPTION OF PROJECT:

Goals and objectives. Goals are to enable non-English-speaking children to achieve proficiency in their mother tongue (Spanish) and in a second language (English), which will permit them to master the school curriculum in both languages, and to cultivate pride in their native language and culture. By achieving these goals, it is hoped that project students will reach grade level in all subject areas and that students who are bilingual will be able to make contributions to both cultures.

Context. Kingsville, in Kleberg County, Texas, is about 36 miles southwest of Corpus Christi and 120 miles east of the border town of Laredo. The population is 28,000 and is somewhat mobile due to a naval base located in the city. Attrition from the program, however, has not been a problem. A celanese plant just outside town employs many people; other employment for residents is with the city, in services, and with local businesses. Kingsville is also the home of Texas A & I University.

There are 6,505 students enrolled in the Kingsville Independent School District, many of whom come from homes in which Spanish is the dominant language. The ethnic breakdown in the district is 54% Mexican-American, 3% black and 43% other. At the elementary level (K-5), about 50% of enrolled children are non-English dominant. The one school that was the nucleus of the program for four years was almost totally Mexican American.

Program description.

Major features of the program are the following:

- use of behavioral objectives and curriculum guides,
- paired teaching in grades 1-5,
- emphasis on language development in both languages,
- systematic assessment of oral language abilities,
- preservice and inservice training, and
- parent and community involvement.

These features will be discussed in appropriate sections below.

Grade level(s), years of operation, size. The program is now in its fifth year of operation. In Year 4, the program served 383 children in grades K-4 in one school. This year, in order to desegregate, the district clustered two grades in each elementary school. There are now three schools involved in the Title VII bilingual education program. These schools constitute one of three elementary level clusters. Kindergarten and first grades are in one school; second and third grades are in another; and fourth and fifth grades are in the third school. In these three schools, there are now almost 500 children involved in the bilingual education program in grades K-5. In addition, one parochial school is providing bilingual instruction for approximately 70 children.

Staffing. Last year, full-time staff members were the director, 14 teachers, 9 teacher aides, and a secretary. There was also a bookkeeper who worked half-time on the project. Some classrooms also have student teachers from Texas A & I University. The evaluator and an assistant were subcontracted to manage the testing and analysis of data. The assistant is also responsible for administering taped morphological and phonological tests to program participants. Several other people at the district level assist the staff of the bilingual program when necessary. These people are the Curriculum Director, Elementary Supervisor, Counselor, and Visiting Teacher, who visits parents at home.

Curricula, time involved, facilities.

Behavioral objectives -- These have been developed for each grade level in the program and are contained in curriculum guides. The guides suggest procedures and materials that can be used to teach the objectives and specify the language(s) in which objectives are to be taught. Content areas taught bilingually are the following: language arts in Spanish and English, social studies (including health, safety, and culture), and mathematics. At the kindergarten level, there are additional objectives relating to school orientation (i.e. awareness of school facilities, respect for school personnel, respect for equipment). The director spot checks teachers during the year to make sure they are teaching to objectives. Teachers may make slight adjustments in objectives to fit needs of the children.

When the guides were written it was decided that certain vocabulary words in Spanish would be used while teaching, especially when referring to elements of the home and school environments. This was done to minimize confusion for teachers and to lay the foundation for a good standard Spanish vocabulary. The procedure was also used to eliminate unfamiliar language found in some of the readers that came from Mexico. If a word was used that was foreign to the children's experience, teachers crossed it out and substituted the word that had been agreed upon as being more relevant. This provides consistency across the program while also developing the children's vocabulary in Spanish.

The objectives in Spanish language arts provide for reading experiences that will enable students to develop decoding skills necessary for the effective use of reading for learning. English language arts objectives are designed to teach Spanish-dominant children proper usage of the English language. Objectives for math and social studies relate closely to the content of the textbooks used throughout the district.

Kindergarten -- All kindergarten children take the Inter-American General Ability Test in Spanish and English. With parents' permission, those who demonstrate a language handicap by scoring below a certain percentage are enrolled in the bilingual education program. The core of the kindergarten curriculum is a locally developed book containing 32 lessons, each of which teaches one sound in the Spanish language. Each lesson includes a high interest, action oriented story about the sound and a variety of reinforcement activities. Stories were adapted from books from Mexico. The grapheme is introduced simultaneously. Lessons also contain questions about the story, vocabulary words using the sound, and a rhyme. To reinforce the sound, children may be asked to think of other words beginning with that sound, or to write the letter on the board, or to find the letter on a chart. Workbooks introduce the letter in cursive form, although children are not formally taught manuscript writing until grade 2. Teachers generally devote one week to each lesson, but there is flexibility to accomodate different learning rates of children. Students then learn how to combine letters to make syllables and syllables to make words, so that by grade 1, most children have acquired the skills for reading.

Kindergarten classes are self-contained, and all kindergarten teachers are bilingual. Instructional time at all grade levels is equal for both languages (50% Spanish and 50% English). However, kindergarten children with very limited abilities in English may be taught primarily in Spanish until their English language development is sufficient to cope with completely bilingual instruction.

Grades 1-5 -- Reading instruction in Spanish begins in grade 1 and continues to expand upon skills taught in the kindergarten program. Once the child has mastered the skills and begins to read, the focus shifts to developing fluency and comprehension. Instruction in English language arts in grade 1 continues to upgrade students' verbal abilities, expand vocabularies, and enhance understanding of the sounds and structure of English. Children who demonstrate sufficient readiness on the Inter-American reading readiness test begin reading in English during the second semester of grade 1. For those whose English speaking vocabulary needs additional development, reading in English is prolonged until they show readiness. By the end of grade 2, all children are reading in Spanish and English. Reading instruction in English adheres closely to the reading series used in the district. The curriculum in social studies and math is also centered around the texts used district-wide. Achievement in these areas is measured by tests provided with the textbooks.

In grades 1-5, classes are integrated so that only part of each class is involved in bilingual education. Students move back and forth between classrooms. Instruction in content areas continues to be equally divided between Spanish and English, but scheduling is such that students in the bilingual program are taught in Spanish by one teacher and in English by another. Teachers are paired at each grade level, and not all teachers at the elementary level are bilingual. Paired teachers plan together to ensure that instruction is coordinated and mutually reinforcing. For instance, if the bilingual teacher is teaching a unit in Spanish on plant reproduction, the other teacher will emphasize the vocabulary and concepts during language arts instruction in English.

Oral language instruction (K-5) -- To develop facility in both languages, oral language drills were developed that teach and reinforce the structure, sounds, and syntax of English and Spanish. They are an integral part of the curriculum of all grades and build on skills acquired in each preceeding grade. Drills are practiced aloud by students for three to five minutes each day. Teachers vary the procedures to avoid boredom. Accompanying drills are a series of morphological and phonological tests that are administered on a pretest/posttest basis. The tests measure students' abilities to make grammatical changes (i.e. knowing plurals for irregular nouns or changing verb tenses), to speak with intonation and stress, to pronounce vowel sounds, etc. Tests also provide for a review of commonly misused speaking patterns often used by Spanish-dominant children and for vocabulary expansion. Each child's pretesting session is taped, and weak areas are tallied. The results are returned to the teacher, who structures much of his or her instruction in language development around these identified weaknesses. Posttesting occurs when teachers feel that students have mastered the drills.

Classroom management techniques -- A pupil profile chart is maintained on each child in the bilingual program. It is a cumulative record of his pretest and posttest scores on all standardized tests, on tests that accompany textbooks, and on locally developed language tests. Grouping patterns are determined by teachers and are a function either of degree of language proficiency or achievement level. Within a grade level, groups have different compositions for each subject area, and they change as often as children's needs change.

Facilities -- The program operates in regular classrooms. No modifications of school facilities are required.

Materials. A summary of the key materials used at all grade levels follows.

In addition to the locally developed storybook used in kindergarten, reading readiness skills are taught with the Language Skills Text from the Key to Reading Series by the Economy Company, the Learning to Think Series by Science Research Associates, Alpha Time with the Huggables by New Dimensions in Education, and a large selection of audiovisual equipment.

Houghton Mifflin's *Preparándose Para Leer* is the core of the curriculum for reading instruction in Spanish in grade 1. Other readers used are *Mis Primeras Letras*, *Felicidad*, and *Victoria*, all published in Mexico. *Ejercicios de Lenguaje* and *Mi Primera Gramática*, also from Mexico, supplement language arts instruction in Spanish. El Nuevo Sembrador, a reading series published in Spain, is the Spanish reading program used in grades 2-5. One of the books is also used in grade 1. Supplementary Spanish readers of varying degrees of difficulty, library books, and audiovisual aids are also utilized. Many of the library books are from Mexico.

Reading in English is taught in all grades with the district-adopted Houghton Mifflin Reading Program. There is close monitoring at the district level of all students' reading achievement in English.

In grades K-3, mathematics is taught in Spanish and English with the Addison-Wesley series, Elementary School Mathematics, and in grades 4-5 with Modern School Mathematice from Houghton Mifflin.

Social studies is taught with *Texto de Unidades Didácticas* (Diploma) from Spain and Lands of Promise published by MacMillan.

The science curriculum uses *La Ciencia en Su Vida* from D.C. Heath and Company and *Science* by Laidlow, a division of Doubleday & Company.

Preservice/inservice training. For three weeks each summer, teachers attend a series of preservice training workshops. They are scheduled by the director and are based on particular needs of both teachers and students. Consultants in many aspects of bilingual education have spoken on the following topics: linguistics, teaching methods in Spanish for speakers of English and Spanish, use of materials and equipment, and team-teaching methodology. Teachers have presented mini-lessons demonstrating different techniques. They also have an opportunity to prepare and develop materials for the coming school year.

During the year, monthly inservice training meetings are held for all members of the bilingual staff. As with the summer workshops, topics are determined by specific needs of teachers and students. These meetings have been used to further orient new teachers to the project and its objectives and to provide more experienced teachers an opportunity to plan additional classroom activities. Policy changes within the program or at the district level are also discussed. Teachers evaluate the usefulness of all training sessions and make suggestions of other subjects they would like to pursue in their training.

A third element of the staff development component is video-taping of teachers while they teach. This is done three times a year by the director. As she observes, she records the types of activities in progress, student reactions to the work, and whether the activities relate to program objectives on a teacher effectiveness analysis form. She also notes physical characteristics of the classroom, such as neatness and displays. She then reviews the videotape with each teacher, praising effective techniques and suggesting alternate approaches for less effective methods. She also points out incorrect language usage. With the teachers' permission, the director has shown some of these tapes at workshops conducted for other bilingual education programs. Teachers are working toward certification in bilingual education, soon to be a requirement in Texas. Certification involves attending a 30-hour workshop given by one of the Texas regional service centers, at least one year teaching experience, and demonstrated proficiency in Spanish. For monolingual English-speaking teachers who are now teaching in the program, this entails receiving 200 hours' instruction in Spanish and passing a language proficiency test.

Parent/community involvement. Participation by parents and community members in Kingsville's bilingual education program and parent-community awareness of its philosophy have always been regarded as prerequisites for program success. Major features of the parent involvement component are an advisory board, a special parent questionnaire, a parent education program, and publicizing the program.

The Advisory Board for the Bilingual Program, organized at the program's inception, is composed of 12 people and meets four times a year. Members of the Board are parents, representatives from community organizations and agencies, and personnel from the district. Parents not on the Board, teachers, the project director, and other school staff generally also attend meetings, which are scheduled at a time when the most people can attend. This is usually at noon. Agendas of the meetings include an explanation of the program's beginnings, its philosophy and goals, demonstrations of classroom activities, planning for the following year, and reviewing proposals. The Board also solicits suggestions from parents about ways to improve the program. This year, because of district-wide clustering of the schools to achieve desegregation, the Advisory Board and Title VII staff has communicated the program's philosophy to the two schools that recently became involved.

A parent questionnaire is distributed at the beginning of each year to assess parents' willingness and availability to participate in activities such as helping with parties, chaparoning or driving for field trips, telling stories in the classroom, or being a parent representative to organizations or out-of-town meetings. Parents are encouraged to volunteer for at least one activity. They are always welcome to visit bilingual classes and are kept informed of special activities by notes sent home with their children.

Parent education has been another vehicle for assuring confidence in the program and enthusiasm for learning. Title VII has sponsored workshops for parents on drug education, health education, and sex education. Parents are informed of other services and opportunities available to them through the district, such as basic adult education classes. Parents have attended workshops in other communities on bilingual education.

Information about the project is disseminated to the community through newspaper articles. The project also tries to present one or two programs for the community each year that stress the culture and traditions of Mexican Americans.

Costs. Budgets and enrollments for the first four years of program operation were as follows:

	Enrollment	Total Budget	Per-pupil cost
Year 1	100	$ 97,800	$978
Year 2	205	148,907	726
Year 3	293	172,169	588
Year 4	383	213,717	558

For Years 1 - 3, the per-pupil cost is computed on the combined total of Title VII and local funds,including teacher salaries. In Year 4, the program began receiving state funds, which are included in the per-pupil cost for that year. This figure is the amount expended for the total educational program for each child in the bilingual education program. The per-pupil cost for the regular district program in Kingsville was $835 in 1972-73 (Year 3 of the bilingual program) and $954 in 1973-74 (Year 4) and includes all expenditures except debt services and capital outlay for equipment with a life expectancy of more than one year.

According to program staff, the reason for the drop in per-pupil cost each year is because the program was contained in one school for four years, and materials were accumulated and shared in that school. For example, most of the major instructional materials purchased in Year 1 for grades K and 1 were still being used in 1973-74, but were shared by more children. The same pattern held as the program expanded vertically to grade 4 by 1973-74.

Major start-up costs incurred in Year 1 were $5,265 for contracted services, including the auditor, the evaluator, and consultants; $5,160 for classroom and office furniture, audiovisual equipment, and nonconsumable instructional materials (textbooks, library books, etc.); $940 for consumable materials; and $2,540 in travel expenses for the administrative, teaching, and consulting staffs. The program also received a one-time planning grant for $5,167. Contracted services continued to be a significant expenditure and averaged about $7,450 for Years 2, 3, and 4. Average travel expenses for the past three years increased slightly to $3,100. Expenses for instructional equipment and materials has increased as the program has increased its enrollment. Costs for furniture, audiovisual equipment, and nonconsumable instructional materials have averaged around $10,000 and about $3,300 for consumable materials and supplies.

EVIDENCE OF EFFECTIVENESS:

Overview. The Title VII Bilingual Education Program in Kingsville, Texas is in its fifth year of operation. The total community is characterized by low socio-economic conditions. Therefore, school achievement tends to be below average at all grade levels.

The evidence of program effectiveness reported in this section is based upon data gathered for the 1973-74 school year. Bilingual classes in kindergarten through grade four that year were located in Colston Elementary School. Control classes at the same grade levels were located in Kleberg Elementary School.[a] Achievement test means for bilingual and for control classrooms in kindergarten through the fifth grade are compared.

Evaluator. The evaluator of the Kingsville bilingual program is Dr. Marlyn Barlow, Professor of Education, Texas A and I University, Kingsville, Texas.

Pupils. Table 1 shows the number of classes and the number of pupils for whom pre- and post-test scores were available on both the Spanish and English versions of the Inter-American Series.

TABLE 1

1973-74 Bilingual and Control Classes and Number of Pupils

Group	Number of Classes K	1	2	3	4	Total	Number of Pupils K	1	2	3	4	Total
Bilingual	3	4	3	3	3	16	48	71	65	77	55	316
Control	2	2	2	2	2	10	41	17	27	28	33	141

During the early history of the programs, attempts to assign classes in Colston to a control group failed because program materials would be shared by the bilingual classroom teachers at Colston. Consequently control classes were assigned at Kleberg Elementary School, a school that matched Colston in terms of socio-economic level and ethnic mix. Both schools contained approximately 97-98 percent Spanish-surnamed pupils.

Attrition is very low within the bilingual program group. The community is stable, so that there is over a 90 percent pupil carryover from year to year.

Measures. The following tests were administered to the pupils in both the program and control groups.

- The SRA Primary Mental Abilities Test was administered only in the first and third grade bilingual classroom in the fall of 1973. Scores obtained were verbal meaning, space relations, number facility, perception, total score.

---

[a] During the first four years, the program was implemented only in the Colston School.

- Both Spanish and English versions of The Inter-American Series: Tests of General Ability were administered to bilingual and control pupils in September 1973 and April 1974. Classroom raw score means are reported and all possible t-test comparisions between classroom means were made.
- The SRA Achievement Series was administered in April 1974 to bilingual and program pupils in the first, second and third grades. The intention was to predict SRA Achievement Series spring scores from the best-weighted linear combination of SRA PMA score and Inter-American Series score obtained the previous fall. These predictions were used for defining individual problem areas so that special attention could be given to pupils with low predicted scores in certain academic achievement areas.

Methods of comparison. Each classroom mean was compared with other classroom means within a grade level. The significance of the mean difference between pairs of classrooms was based upon a t-test statistic. For the SRA Primary Ability Test comparisons, mean differences were tested between all pairs of the four first grade bilingual classrooms and between all pairs of the three third grade bilingual classrooms. For each of the five PMA scores, there are six possible contrasts in the third grade and 10 possible contrasts in the first grade. Significant differences among the classroom groups were revealed at both grade levels.

The major comparison was based on mean gain scores on the Inter-American Series. At each grade level, all possible classroom mean gain contrasts involving a bilingual and a control classroom are tested by means of t-tests. Thus, there is no overall test of total bilingual versus total control group mean gain at a grade level, but instead, what is reported is the number of possible contrasts involving a bilingual and a control classroom and the number of these that significantly favor either the bilingual or control classroom.

No data in the SRA Achievement Tests are reported.

Results. Mean classroom comparisons on the SRA-PMA all involved bilingual pupils, and are not tabled in this section. It is important, however, to note that there were significant mean classroom differences on all of the PMA tests.

Table 2 classifies all possible mean gain comparisons involving a control and a bilingual classroom at each grade level according to whether the comparison was significant or not and whether the mean differences favored either the bilingual or the control classroom.

Of the 70 possible bilingual-control classroom comparisons across all five grade levels, 25 showed significant mean gain differences on the Inter-American Series in favor of the bilingual classroom, and five showed mean gain significance in favor of the control classroom.

On the Spanish version of the Inter-American Series, 24 mean gain comparisons favored the bilingual group and 17 of these were significant at the alpha = .05 level. Eleven comparisons favored the control group and none of these were significant. Thus in terms of the acquisition of Spanish language

TABLE 2

Classroom T-Test Comparison Survey: 1973-74 Kingsville

Results	K		1		2		3		4		Total	
	B[a]	C[b]	B	C	B	C	B	C	B	C	B	C
English:												
Significant[c]	6	0	1	2	0	0	0	3	1	0	8	5
Not Significant	0	0	3	2	4	2	2	1	5	3	14	8
TOTAL	6	0	4	4	4	2	2	4	6	3	22	13
Spanish:												
Significant	6	0	2	0	4	0	2	0	3	0	17	0
Not Significant	0	0	2	4	1	1	1	3	3	3	7	11
TOTAL	6	0	4	4	5	1	3	3	6	3	24	11

[a] B = Bilingual mean higher than control
[b] C = Control mean high than bilingual
[c] = Significant at the alpha = .05 level

skills, the bilingual classes outgained the control classes at all grade levels. Nearly half (17 out of 35) possible bilingual-control contrasts were significant and all favored the bilingual classroom.

On the English version of the Inter-American Series, there were 13 significant and 23 insignificant mean gain differences between bilingual and control classes. Of the 13 significant contrasts, eight favored the bilingual class and five favored the control; six of the eight comparisons favoring the bilingual class were at the kindergarten level; two of the seven significant contrasts for grades 1 through 4 favored the bilingual classes. While the evidence is not as strongly favorable on the English as on the Spanish versions of the test, the bilingual classes nevertheless did outgain the control classes in kindergarten.

APPENDIX E

OUTLINE FOR
BILINGUAL EDUCATION
PROGRAM DESCRIPTIONS

BILINGUAL PROGRAM DESCRIPTION OUTLINE

I. Program Overview (500-800 word summary)

Identification information
- program title
- location name
- year started

Background information
- historical development
- context
- needs assessment

Objectives and procedures
- bilingual/bicultural philosophy
- major objectives and rationale
- primary program features and rationale
- key instructional strategies
- time devoted to instruction in both languages
- use of language in content area instruction

Participants
- age, grade level, sex, number
- demographic background
- language dominance
- special characteristics
- language baseline competencies
- selection criteria (qualifications)
- selection procedures (test, referral, screening panel)

Personnel
- categories and number
- qualifications and training

II. Program Development

Relation of program to target population and school system
- target population size and composition
- demographic and cultural factors
- geographical area served, physical size, number of schools by grade level
- integration of program in system and with other programs

History and needs assessment
- origins and philosophy
- needs assessment focus, methods, and priority selection
- impact on program development

Bilingual/bicultural program objectives and rationale
- objectives with respect to language-dominant subgroups
- general strategy for measuring attainment of objectives
- expected outcomes

Planning the program (prior to start)
- key persons/roles in major planning tasks
- major steps, e.g.
    - enlisting support
    - identifying resources
    - obtaining funds
    - developing/adapting curriculum materials
    - assembling and training staff
    - selecting participants

Changes/growth in original program[1]
- pattern of growth and development
- major changes, if any, and rationales for them
- major problems, if any, and strategies for solving them

III. Staffing and Management

Staff types, numbers, and chain of command
- job descriptions and roles
- time devoted to program, length of service in program
- qualifications, certification, experience in bilingual education
- competencies in second languages

Recruitment of staff
- sources
- bilingualism requirements and experience
- selection strategies
- strategies for maintaining a stable, motivated staff

Preservice and inservice training
- objectives
- relationship to wider bilingual/bicultural program objectives
- key persons/roles in conducting training
- extent, schedule, activities, methods of training
- assessment of training program

Management strategies
- how chain of command is implemented on daily basis
- program administrator's autonomy and authority at district and school levels
- key management strategies and rationales

IV. Instruction

Language of instruction
- choice
- rationale
- extent of use in instruction
- attitude of teachers and children to language of instruction

---

[1] Writers were free to discuss changes the program had undergone in each chapter, as appropriate.

Instructional activities by grade level and by language/content areas
- relation of methodology to objectives and philosophy of program
- bilingual teaching techniques
- other teaching techniques
- cultural component of the curriculum
- motivational/reward techniques
- feedback to students and parents
- classroom management
- classroom climate

Grouping
- objectives and purposes
- criteria and patterns
- frequency of grouping children for daily instruction
- frequency of regrouping

Diagnosis and assessment of participants
- specific competencies measured
- measurement procedures, instruments, frequency, records, reporting
- use of diagnostic/assessment records for decisions
- other observational procedures

Typical schedule/timeline by grade level
- time devoted to dominant and second language instruction in each subject
- integration of language subgroups for scheduled activities

Physical layout of instructional facilities
- features of plant that affect instruction
- features of classroom that affect instruction

Key materials and equipment
- list of key items and quantity required
- procedures/rationales for selection, adaptation, development, and use
- cultural reference/language of key items of materials

V. Parent Involvement and Community Awareness

Rationale and purposes
- relation to program philosophy and objectives
- intended outcomes

Historical perspective
- trends in participation
- changes in scope or emphasis

Specific nature of involvement
- key groups and their impact on the program
- school- and class-based activities
- home-based program activities
- extra-school involvement in program planning and review

Measures of involvement

Results

VI. Costs

Sources and level of funding
- Federal funds
- State funds
- Local funds
- Other funds

Analysis of program-supplied cost information
- Start-up costs by number of participants
    - major cost categories
    - one-time costs
- Continuation costs by number of participants for succeeding years
    - major cost categories
    - consumable items, reusable items
    - cost trends
- Per-pupil costs
    - basis for computing
    - state or local per-pupil allocation to which added on, if applicable
    - trend in per-pupil costs
    - compared to per-pupil cost of regular program

Guidance for developing a replication budget
- Further interpretation and comments re cost data
- Budget options

VII. Evaluation

Design
Measures
Results
Interpretation

VIII. Sources for Further Information

Program director
Evaluation director
Materials and equipment sources
References